Event histories in the Netherlands Fertility and Family Survey 1998. A technical report

Matsuo, Hideko; Willekens, Frans

Event histories in the Netherlands Fertility and Family Survey 1998

A technical report

Hideko Matsuo and Frans Willekens

Population Research Centre
Research Report 03-1, February 2003

Population Research Centre

PO Box 800
9700 AV Groningen
The Netherlands
Tel +31 50 363 3898
Fax +31 50 363 3901
e-mail: PRC@frw.rug.nl
www.rug.nl/prc

Event histories in the Netherlands Fertility and Family Survey 1998

A technical report

Hideko Matsuo and Frans Willekens

Population Research Centre
Research Report 03-1, February 2003

Acknowledgements.
We like to thank the students in the MSc in Population Studies programme, who made extensive use of the Netherlands Fertility and Family Surveys 1993 and 1998 as part of the coursework for 'Life history data analysis' and 'Multistate Demography'. Special thanks goes to drs. José Dias, who developed the procedure used in this paper to infer the month of interview. The contribution of drs. Karen Haandrikman in the preparation of the English version of the original Public Use Data File is greatly appreciated. The draft report was submitted to Statistics Netherlands. We are grateful to drs. Arie de Graaf of Statistics Netherlands (CBS) for his careful reading, his comments and suggestions. The report also benefited from the extensive comments by dr. Sabu Padmadas, drs. José Dias and drs. Sarbani Banerjee.

EDITORIAL

Research Reports provide a formal record of research conducted by members of the Population Research Center. They offer a level of detail or technicality that is often required but is usually not suitable for a journal publication. Research reports are externally reviewed before publication.

This is the first research report. It approaches the information collected in the Netherlands Fertility and Family Survey 1998 (OG98) from an event history perspective. Increasingly, individual data on life events are made available for public use. Event history models and multi-state models are the dominant analytic instruments today. Most event history models consider one event at a time while multi-state models consider sequences of states. In order to get most from the data, the sequence and timing of events must be determined accurately. In case of anomalies in the data, specific assumptions about sequence and timing of events must be made. It is however not always common to document and discuss the procedures in great detail. This report presents a detailed assessment of the event histories reported in the OG98, provides examples of abnormality and inconsistencies in the data, and identifies information that is missing but should be available for event history analysis and multi-state analysis.

This report shall however be useful not only to readers using OG98, but also to others who embark on the analysis of similar individual data. It is a unique type of research report that highlights some areas that are often not considered. The report shows that any type of quantitative analysis should pay sufficient time and energy before embarking on any sophisticated type of analysis. Basic comes first.

I wish you a pleasant reading.

Karen Haandrikman
Editor

ABSTRACT

The Fertility and Family Survey (FFS) [Onderzoek Gezinsvorming (OG)] is designed by Statistics Netherlands (CBS) to supply basic data to formulate hypotheses for the Population and Household Forecasts. This research report describes the conversion of the Public Use Data File of the OG 1998 (female respondents, individual file) into an event history data structure that facilitates event history analysis. The main requirement is the ordering of events and the definition of events in terms of origin state, destination state, and date of occurrence. All the dates are recoded in century month codes (CMC). We consider more than 20 events related to leaving the parental home, marriage, cohabitation, and childbearing. For each respondent, the OG98 reports up to three marriages and up to six cohabitations. Investigation of the sequences of events reveals a few measurement problems and inconsistencies. They are identified and removed using additional information provided by the survey or relying on explicit assumptions. The month of interview is estimated since the information is omitted from the Public Use File. The information is necessary to estimate transition rates in the presence of censoring. The report provides the SPSS syntax that converts the Public Use File into an event history data structure.

Keywords • Event history analysis • Netherlands Fertility and Family Survey • Life course • Transition data • State space

CONTENTS

LIST OF TABLES

LIST OF FIGURES

TEXT BOXES

1. INTRODUCTION

During the recent decades, the life course paradigm emerged as a dominant perspective in demography and other social sciences (Giele and Elder, 1998). Two reasons may be provided. One reason for the emergence of the new paradigm is theory-driven and the other is data-driven. In order to comprehend the reasons why people marry, cohabit, have children, divorce or migrate, the demographic life transitions must be viewed against the background of the experiences people accumulate and the expectations they have about how these transitions will affect their lives (McDonald, 1996, p. 385). The experiences and expectations of individuals depend not only on personal factors but also on the social and historical context. The life course perspective differs from traditional perspectives in a number of ways. First, it emphasizes that the personal, social and historical factors operate across all stages of the life span. Second, it asserts that experiences during critical or vulnerable periods of life have a greater effect than experiences during other periods. Third, it distinguishes immediate effects from medium- and long-term effects. Fourth, it looks for life patterns that emerge because effects span across time and most people strive for behavioral consistency. The timing of transitions in life is often measured in more than one time scale to highlight the distinction between innate factors and contextual factors. Individual time and age and historical time are two dimensions that are commonly used. Both contemporary factors and historical factors operate in historical time. The first set is known as period effects and the second as cohort effects. The study of the interaction between individual life and history (social change) goes back to Ryder (1965) and his seminal paper *The cohort as a concept in the study of social change*. Recently, Elder (1999, p. 5) pointed to Ryder's work as a 'point of departure for understanding the interaction between social change and the life patterns of birth cohorts'.

The emergence of the life course paradigm is also linked to the increased availability of longitudinal data. The rationale of longitudinal data and longitudinal research has long been established (see e.g. Baltes and Nesselroade, 1979). Longitudinal data are collected to examine behavioral changes and to determine the part of the change that may be attributed to personal experiences or interventions (e.g. policy measures). The effect of an experience or an intervention is not always immediate and it may be enhanced or inhibited by intervening factors or confounding factors. Repeated measurements reveal the dynamics (for a discussion, see e.g. Willekens,

2001). They are particularly helpful for making causal inferences and causal statements (Blossfeld and Rohwer, 1995, pp. 20ff.). Although the virtues of longitudinal data are beyond question, it is often not feasible or it may be too costly to follow respondents for an extended period of time. A research design that collects life histories retrospectively is a viable alternative, though it is vulnerable to recall biases of human memory and to selection biases since only survivors report their life history (for a discussion, see e.g. Scott and Alwin, 1998).

Life histories may be described by recording the attributes of a respondent at consecutive points in time or by recording the attributes at an initial time and the events that change the attributes. The first approach has been referred to as the status-based approach, the second as the event-based approach (Willekens, 2001). An example of a status-based approach that combines retrospective measurement and prospective panel design is the life history calendar (Freedman et al., 1988; Belli et al., 2001; Khatun and Willekens, 2001). The calendar typically documents, for every month of an extended period of time, one or several attributes of the respondents. The attributes are represented by qualitative or discrete variables. The values of the variables change with age and/or with time. The calendar captures that change. The second approach focuses on events. Instead of recording the attributes at several points in time, *changes* in attributes are recorded. The design is an event-oriented design and the data are referred to as event history data (Blossfeld and Rohwer, 1995, p. 17). The observation of events and event sequences (event histories) provides an alternative to the observation of sequences of attributes or states (for a detailed treatment, see Tuma and Hannan, 1984, pp. 18ff; see also Blossfeld and Rohwer, 1995, pp. 17 and 33ff.). The aim of this report is to document the conversion of data that are collected retrospectively in the absence of a calendar design or an event-oriented design into an event history data structure. The event history data structure facilitates event history analysis and multistate life table analysis.

Our principal objective is to convert the Public Use Data File of the Netherlands Fertility and Family Survey 1998 into an event history data structure to subsequently facilitate event history analysis and multistate life table analysis. The Netherlands Fertility and Family Survey (FFS) 1998 [Onderzoek Gezinsvorming (OG98)] is a rich source of life history data. It was not designed to study sequences of events, however. It was designed by Statistics Netherlands (CBS) to supply basic data to formulate hypotheses for the Population and Household Forecasts. The conversion of raw data into an

event history data structure primarily requires data to be arranged as event sequences and accurate measurement of the timing of consecutive events in the life course. The emphasis on the sequence and timing of events may reveal several inconsistencies in the data that remain hidden otherwise. Particular sequences of events may not be possible (e.g. second child is born before first child) or plausible (e.g. marriage before leaving the parental home). Events may be missing (e.g. second marriage is reported while information on dissolution of first marriage is missing). Our investigation reveals that there are such inconsistencies in the OG98. The inconsistencies may be real or may represent behaviour that deviates from plausible behaviour. Several apparent inconsistencies are associated with deviant behaviour. Since these inconsistencies might signal early stages of social change, they are studied with great care. When inconsistencies in the data are clearly related to errors in recording or coding and the correct values can be derived from other information in the data, corrections are made. All the corrections are reported in this paper.

The collection and analysis of life history data has been a subject in many studies (e.g. Yamaguchi, 1991; Courgeau and Lelièvre, 1992; Blossfeld and Rohwer, 1995; Vermunt, 1997; Mills, 1999, 2000; and several chapters in Giele and Elder, 1998). For a specific discussion of data collection and data structure, the reader is referred to Blossfeld and Rohwer (1995, Chapter 2) and Brueckner and Ulrich Mayer (1998). Most studies that apply techniques of event history modeling do not consider event histories, i.e. sequences of events. They consider one event at a time and apply techniques of survival analysis. Even Blossfeld and Rohwer (1995) address one event at a time and single episodes although some reference is made to multiple episodes. For instance the sequence consisting of a job episode followed by an episode of unemployment is not considered. As a result, some anomalies in the data remain undetected. As an illustration, the data supplied with the book (file rrdat.1) include five respondents who start a new job episode before the previous job is completed (cases 17, 110, 113, 140, 150). We report several similar anomalies in the data of the Fertility and Family Survey. The conversion of a data structure into an event history data structure is a tedious process. We report all events that are seemingly inconsistent. The inconsistency of a reported sequence of events may be attributed to non-response, recall problems or inaccurate coding. The same sequence may however be real too. For instance, we found that some women married or started cohabiting before they left their parental home. It is impossible to determine whether this sequence of events is real or a limitation of the data.

The questionnaire design, which distinguishes between leaving home for reason of marriage or cohabitation, and for other reasons, points in the direction of a real inconsistency.

This paper begins with a description of the OG98. We focus mainly on the survey design. Analysis of results is beyond the scope of this report. Results of the survey were presented by De Graaf and Steenhof (1999) and Garssen et al. (2001), among others. Section 2 discusses the main features of the OG98 including the sample and the questionnaire. Section 3 presents a conceptual framework that has shown to be very useful in life history analysis. It is the description of the life course in terms of states occupied at various ages and the transitions between states. The state space perspective provides the overall framework guiding the conversion of data into an event history data structure. State variables measure the attributes of people at various points in time. Transitions change the values of state variables. The timing of transitions is translated into a uniform Century Month Code (CMC). That transformation facilitates the study of episodes and transitions. The description of the state space and the timing of transitions constitute the subject of Sections 3 and 4. Once the state space is defined and the timing of transitions is determined, the individual life paths are fully documented. The investigation of event sequences reveals a number of measurement problems and inconsistencies. Inconsistencies may be real or they may be the consequence of the definition of the state space. Measurement problems and inconsistencies are the subject of Section 5. Section 6 fills an important gap left by the OG98; namely, the estimation of the month of interview. The Public Use File does not include the month of interview. Since that information is necessary for the analysis of event histories, the missing data are imputed. The conversion of the Public Use File of OG98 into an event history data structure involves several steps. They are documented in the report. The conversion is implemented using SPSS syntax. The syntax is included in the report and can be downloaded from the website of the Population Research Centre of the University of Groningen. The syntax and the variables that are created during the conversion process are described in Section 7. Section 8 concludes the report and formulates a number of recommendations for future fertility and family surveys.

2. THE NETHERLANDS FERTILITY AND FAMILY SURVEY 1998 (ONDERZOEK GEZINSVORMING 1998)

The sample

OG98 collected information on 5,450 women and 4,717 men, born in the period 1945-79 and residing in the Netherlands. The respondents were 18 to 52 at time of interview. As of 1st January 1998, 4.2 million men and 4 million women in the Netherlands were born in the period 1945-79. They were 18 to 52 years at time of survey. This section is based on De Graaf and Steenhof (1999, pp. 35-36).

The survey was conducted between February and May 1998 by the Statistics Netherlands (CBS). The sample frame consisted of the Municipal Population Administration (Gemeentelijke Bevolkingsadministratie; GBA). The GBA is the main source of statistical information on the population of the Netherlands. The random sample survey was done in two steps. In the first step 262 municipalities were selected from a total of 572 municipalities. The GBA of the selected municipalities was used to randomly select 14,000 addresses and subsequently men and women born in the period of 1945-79. The drawing of the random sample was done taking into account several conditions (for details refer to De Graaf and Steenhof, 1999, p.36). Eventually, 5450 women and 4717 men were interviewed by means of structured questionnaires.

The questionnaire

The questionnaire solicits extensive information on cohabitations and marriages; living arrangements; intentions and behavioral outcomes of family formations; and individual background variables. The questionnaire is divided into 10 sub-sections, some of which are detailed while others are not. The sub-sections are:

1. Residential preference (‘Woonwens’)
2. Partnership and partnership formation (‘Relatievorming’)
3. Children: females (‘Inventarisatie kinderen (vrouwen)’)
4. Children: males (‘Inventarisatie kinderen (mannen)’)
5. Societal position and time allocation (‘Maatschappelijke positie / tijdbesteding’)
6. Work and children (‘ Werken en kinderen’)

7. Fertility and birth expectations (‘Vruchtbaarheid en toekomstverwachtingen kinderen’)
8. Attitudes (‘Attituden’)
9. Background (‘Achtergrond’)
10. End of interview (‘Besluit interview’)

The fieldwork

Electronic questionnaires were used in face-to-face interviews. These questionnaires were run by a computer programme called BLAISE. Briefings were held in advance to explain the objective of the survey and to discuss potential problems related to the interview. The interviewee was notified about the interview by letter. To give the interviewee ample opportunity to be present, a minimum of three visits was made. When the respondent agreed to the interview and cooperated in filling in the entire questionnaire, a telephone card equivalent to NLG 2.50 (a little over EUR 1.00) was given as a token of appreciation.

Weighting

Care was taken to ensure that the sample population in the OG98 was representative of the population of the Netherlands. Weighting was applied at the personal level to increase the comparability. The weighting procedure was carried out separately for men and women. Background variables used in the weighting procedure were: year of birth, marital status, position in the household, country of birth, number of inhabitants per municipality. For women, the number of live-born children was included as a weighting variable.

Non-response

Unfortunately, the response rate of the OG98 has not been documented. In the Netherlands, the public willingness to participate in surveys is lower than in most other countries (De Graaf and Lodewijckx, 2000). In the Fertility and Family Survey of 1993 [OG93], the response rate at household level was 48.5 percent and the response rate at the personal level was 90.3 percent (De Graaf and Lodewijckx, 2000, p. 137). The response rate of the OG98 may be assumed to be close to that figure. As a consequence, the results of the OG98

must be considered with caution since it may not be fully representative of the population in the Netherlands.

Data availability

The Scientific Statistical Agency (Wetenschappelijk Statistisch Agentschap [WSA]) of the Netherlands Organization for Scientific Research (Nederlandse Organisatie voor Wetenschappelijk Onderzoek [NWO]) has made available a public use file of OG98. For information, see the website of WSA: http://wsa.magw.nl/index_uk.htm. The data are distributed in two SPSS files. The file BOAV98.SAV contains the data for females and the file BOAM98.SAV contains the data for males. The labels of the variables are in Dutch. English versions of the labels of the variables were prepared by drs. Karen Haandrikman of the PRC, University of Groningen, in cooperation with the first author (Hideko Matsuo).

3. MULTISTATE REPRESENTATION OF LIFE HISTORIES: THE STATE SPACE

In life history analysis, the life course is segmented into domains of life. Family, work, living arrangement, education, residence, and health are such domains. The domains exist in parallel and generally interact. Each domain may be divided further into discrete states of existence. For instance, the following living arrangements may be distinguished: living alone, couple without children, couple with children, and one-parent household. In this example, 'living arrangement' is a variable and the types of living arrangements are the values of the variable. The variable is a categorical variable and the values are the categories. The status of an individual in the labour force may consist of three categories (employed, unemployed and out of the labour force). Each category may be divided further. To characterize a domain of life, a combination of state variables may be used, resulting in a composite variable. In this report, we make extensive use of composite variables. For instance, in family demography, two attributes are used to characterize a state of existence: marital status and living arrangement. If marital status consists of two categories (not married, married) and the living arrangement of three categories (living at the parental home, living alone, living with someone), the total number of categories is six. In order to reduce the number of categories, some may be omitted. Although the identification of relevant states of existence enhances life course analysis, at the same time it also restricts the analysis.

The set of possible states a person can occupy is known as the *state space*. The state space characterizes a domain of life (or several domains of life). Its specification is dependent on the research question, i.e. substantive considerations. The states are mutually exclusive, meaning that a person can occupy only one state at a time. The states are also exhaustive, meaning that at any moment in time, a person in the (sample) population must be in one of the states. The state space identifies not only the states, but also the possible *transitions* between states. A transition is characterized by state of origin, state of destination, and timing of transition. Event history data should be organized in a way that enhances the specification of different state spaces and the associated transitions. We follow the strategy described by Blossfeld and Rohwer (1995, Chapter 2). It involves the definition of the state space and the transitions, and the representation of the time at transition in a convenient way through the Century Month Code (CMC).

The life course starts at birth and ends at death. Except for rare cases, the entire life course cannot be documented or observed. Instead, segments of the life course are observed, e.g. periods of one or ten years. Retrospective surveys generally record the relevant information from birth until the time of survey. In prospective surveys, subjects enter observation at a given age or point in time and leave observation upon occurrence of an event (the event of interest or a competing event) or upon truncation of the observation (e.g at interview). Although the events of interest take place throughout the life course, only events that occur in the *observation window* are recorded. Thus, the starting time of observation and the ending time must be known to permit event history analysis.

The specification of the state space is closely linked to the research question. Since we do not have a specific research question, we present three specifications of the state space, by way of illustration. The first is used to study the marital career of women. Here we are not interested in whether a state is occupied for the first or second time. The second state space describes a fertility career and is commonly used in fertility analysis. It distinguishes children by birth order. The third represents the pathway to the first birth. Since the data are from the Netherlands Family and Fertility Survey, mortality is excluded. The definition of the state spaces also determines the possible sequences of state transitions and the possible pathways to the first child.

a. The first state space categorizes respondents by marital status. Although the order of marriage and marriage dissolution may be considered, the example does not distinguish between first marriage and second marriage, first divorce and second divorce, etc. The states are:

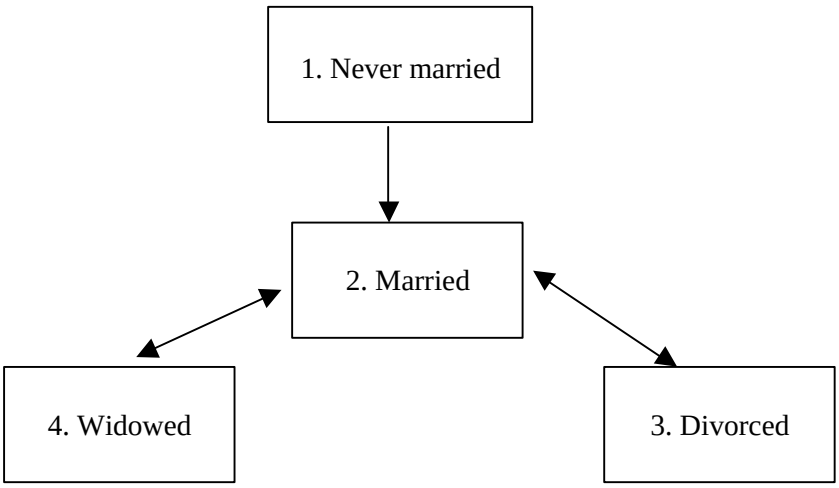
1. Never Married
2. Married
3. Divorced
4. Widowed

The first state ‘Never Married’ is entered at birth. One can leave this state but not enter it again. The other states are transient states, which can be entered and left repeatedly. The state space influences the types of questions that may be answered. For instance, the omission of the order of marriage implies that the behaviour of women in a second marriage cannot be differentiated from the behaviour in the first marriage. The state space is also inadequate to determine the proportion of women of a given age that is in a second

marriage. To answer this question, we need a separate state for each marriage order. It makes the state space more realistic but increases the number of states. The specification of the state space involves a compromise between realism and parsimony. This state space determines the transitions that are theoretically possible. Not all transitions are feasible, e.g. a transition from divorced to widowed. Figure 1 is a representation of the state space and the possible transitions.

This state space is often used to describe the marital career. It is the basis of the marital status life table. At birth, children are in the never married state (state 1). As they get older, they leave that state and marry. The marriage may be dissolved; the reason for marriage dissolution is either divorce or widowhood. The possibility of remarriage and dissolution of the second marriage is indicated by the arrows. However in this model, no distinction is made between first and second marriage.

Figure 1. Schematic representation of four-state multi-state model based on marital status

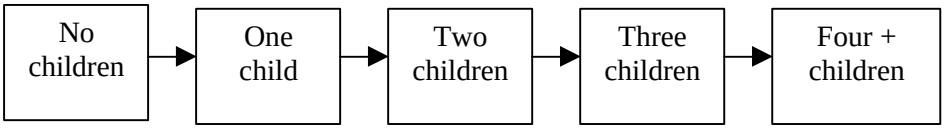


b. Another state space, which is used to describe the fertility career of a woman, is expressed in terms of number of children born. The state space is:

1. No children
2. One child
3. Two children

4. Three children
5. Four and more children

Figure 2. Schematic representation of fertility career



c. A third state space that can be investigated using the OG98 describes the pathways to the first child, i.e. the sequence of states a woman goes through before she gets her first child. The states occupied following the birth of the child are beyond the scope of this investigation. Given this research design, Figure 3 presents the state space and the associated transitions. The path starts with the state of living at the parental home. We assume that the parental home may be left only once, although in reality persons may leave the parental home and return later at least for some time. The respondent may leave home for one of three reasons. The first is independence, which is manifested by leaving home to live alone. The second and third reasons involve union formation through marriage (second reason) or cohabitation (third reason). Childbearing may occur in any of the states. The states are:

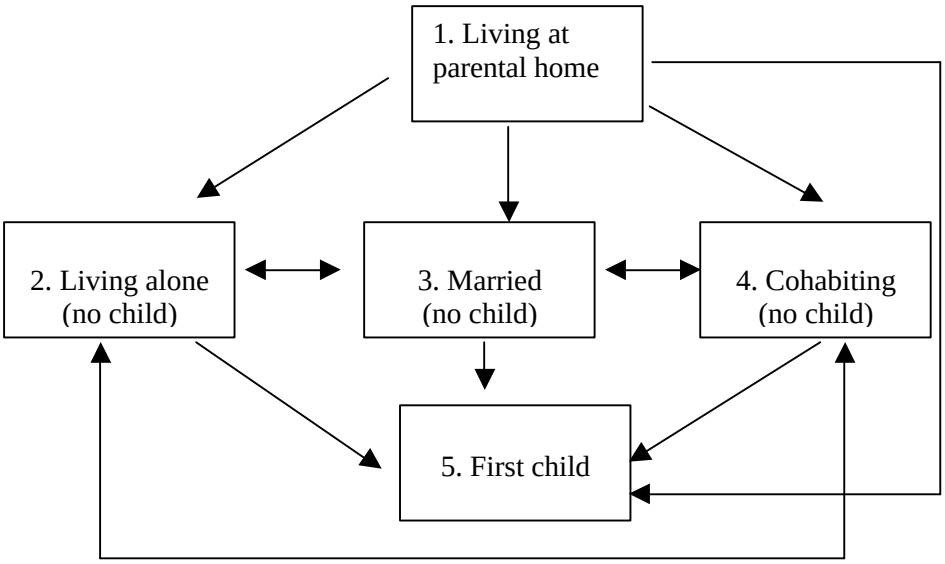
1. Living at parental home
2. Living alone (independently)
3. Married
4. Cohabiting
5. First child

The state space is determined by a composite variable that combines three domains of life. The first domain of life is the living arrangement with three possibilities: living at the parental home, living alone, and living with someone. The second domain of life is the marital status: not married or married. The third domain is motherhood (fertility). The three state spaces are combined into a single state space and some combinations of states are excluded (e.g. cohabitating at the parental home, married while living at the parental home).

The specification of the state space determines the sequence of states and events that can be studied. In the accompanying example, a married woman

may start cohabitation upon marriage dissolution. She may start living alone instead but she may not move back to the parental home. Some living arrangements, such as Living Apart Together (LAT) (commuting marriage), are not considered in the state space and can therefore not be studied. To include that arrangement a distinction must be made between partnership status (union status) and residence status, and the timing of the transitions between the states should be known. The focus on pathways to first birth implies that the transitions that occur after the birth of the child are not considered in the analysis. The birth of a child implies an entry into the final state, which is an absorbing state.

Figure 3. Schematic representation of pathways to first child



The OG98 reveals some uncommon living arrangements. For instance, some married women do not live with their husband; some live alone and some live with another partner. These living arrangements are not considered in this report since we lack information and the state space is too restrictive. To capture these living arrangements, the state space needs to be extended.

4. TIMING OF TRANSITIONS

Throughout the report, a transition is characterized by three attributes: the state of origin, the state of destination, and the timing of the transition. A fourth attribute, the reason for the transition, is not considered in this report, except when needed to determine the destination. The origin and destination of a transition depend on the state space. A state that is not included in the state space can of course not be selected as a state of origin or a state of destination.

In this section, we determine for each transition the timing in terms of the Century Month Code (CMC). The state of origin and the state of destination is the subject of the next section. We consider more than 20 transitions or events that are recorded in the OG98. The events are listed in Table 1. For each event, the OG98 recorded the year and month of occurrence. The variable name in columns 2 and 3 are the same as the ones used in the Public Use File. The variable names for the dates of the events in CMC are shown in column 3. A distinction is made between original dates, i.e. dates not adjusted for inconsistency in the dates and/or sequence of events, and final dates that are adjusted for inconsistencies. The original dates are derived directly from the OG98 and are preceded by O. The final dates are preceded by F.

To facilitate the study of sequences of events and the calculation of intervals between events, the dates of events (years and months) are converted into CMCs. The coding scheme measures the date of the event in months since the beginning of 1900 (see e.g. Blossfeld and Rohwer, 1995, p. 39). An event that occurs in January 1900 occurs in month 1 (CMC = 1) and an event that occurs in March 1960 occurs in CMC 723. The CMCs are calculated as follows:

$$CMC = (\text{Year of event} - 1900)*12 + \text{Month}.$$

The date of event in CMC may be converted into calendar year and month by using the following expressions. The year in which the event occurs is

$$\text{Year} = \text{TRUNC}(CMC/12)+1900$$

and the month is

$$\text{Month} = CMC - (\text{Year}-1900)*12$$

Table 1. Demographic events and corresponding variable names

Event	Year	Month	CMC Original	CMC Final (if different)
(1)	(2)	(3)	(4)	(6)
Birth of respondent	GEBJR_OP	GEBMD_OP	Cmcb_op	
CMC interview (imputed)			Cmcint_o	Cmcint
Cohabitation before current marriage	JRTAL	MND	Cmcocm	
Current marriage	JRTAL2	MND2	Cmccma	
Cohabitation before first marriage	JRTAL3	MND3	Cmcoma1	
First marriage	JRTAL4	MND4	Cmc1ma	Cmcma1
Another estimation of 1st marriage				Cmcma1_1
End first marriage	JRTAL5	MND5	Cmcma1en	Cmce1ma
Reason end 1st marriage		Based on VRG6	Whymd1	Whymd1
Separation in 1st marriage		Based on VRG25		CMCma1se
Cohabitation before second marriage	JRTAL6	MND6	Cmcoma2	
Second marriage	JRTAL7	MND7	Cmc2ma	Cmcma2
End second marriage	JRTAL8	MND8	Cmcma2en	Cmce2ma
Reason end 2nd marriage		Based on VRG10	Whymd2	Whymd2
Separation in 2nd marriage		Based on VRG26		Cmcma2se
Reason husband not in household			Redniet	Redniet
Husband left	JRTAL17	MND17	Cmchleft	
Was this your first marriage			Eersthuw	Eersthuw
Reported number of marriages			Aantkeer	Aantkeer
Serial number of current marriage			ncurma	ncurma
Number of marriages a woman had			Nmarriag	Nmarriag

Table 1, continued

Event	Year	Month	CMC Original	CMC Final (if different)
(1)	(2)	(3)	(4)	(6)
Current cohabitation	JRTAL9	MND9	Cmcurco	
First cohabitation	JRTAL10	MND10	Cmc1co	Cmcco1
End first cohabitation	JRTAL11	MND11	Cmco1en	Cmce1co
Second cohabitation	JRTAL12	MND12	Cmc2co	Cmcco2
End second cohabitation	JRTAL13	MND13	Cmco2en	Cmce2co
Third cohabitation	JRTAL14	MND14	Cmc3co	Cmcco3
End third cohabitation	JRTAL15	MND15	Cmco3en	Cmce3co
Fourth cohabitation				Cmcco4
End fourth cohabitation			Cmce4en*	
Fifth cohabitation				Cmcco5
Which cohab is 1st cohabitation			F_1co	F_1co
Which cohab is 2nd cohabitation			F_2co	F_2co
Which cohab is 3rd cohabitation			F_3co	F_3co
Which cohab is 4th cohabitation			F_4co	F_4co
Which cohab is 5th cohabitation			F_5co	F_5co
Number of cohabitations a woman had			Ncohabit	Ncohabit
Cohabitation contract ?			Contract	Contract
Cohabitation contract	TYDSTJ	TYDSTM	Cmccoco	
Number of unions a woman had			Nunions	Nunions
First union			Cmfunion	Cmfunion
Leaving home (excluding cohabitation. and marriage)	JRTAL16	MND16	Cmclh	

Table 1, continued

Event	Year	Month	CMC Original	CMC Final (if different)
(1)	(2)	(3)	(4)	(6)
Leaving home (including cohab. and marriage)				Cmcleave
Reason for leaving home			Leavewhy	Leavewhy
Ever had children			OOITLEV	OOITLEV
Number of children born alive			AANTLEV	AANTLEV
Birth oldest child	GEBJR_K1	GEBMD_K1	Cmc_k1	
Birth second child	GEBJR_K2	GEBMD_K2	Cmc_k2	
Birth third child	GEBJR_K3	GEBMD_K3	Cmc_k3	
Birth fourth child	GEBJR_K4	GEBMD_K4	Cmc_k4	
Birth youngest child	GEBJR_K5	GEBMD_K5	Cmc_k5	
Ever had adopted/foster/step children			HFTSTIEF	HFTSTIEF
Type: adopted, foster or stepchild by birth order			AARD, AARD2, AARD3	AARD, AARD2, AARD3
Birth oldest step, foster or adopted child	GEBJR_S1	GEBMD_S1	Cmc_s1	
Birth second step, foster or adopted child	GEBJR_S2	GEBMD_S2	Cmc_s2	
Birth third step, foster or adopted child	GEBJR_S3	GEBMD_S3	Cmc_s3	

* This variable is not used. Respondent ID 171 ended a fourth cohabitation and started a fifth. The CMC at the end of the fourth cohabitation was not reported however.

In order to assign a CMC, the year and the month of occurrence must be known. The first event is birth of the respondent. The year of birth is represented by the variable GEBJR_OP and the month of birth is denoted by GEBMD_OP. For most other events, the year is denoted by the variable JRTAL* where * represents a number. In some cases, the year of occurrence is missing (not reported) although the respondent indicated that the event did

occur. In the Public Use File, missing values are represented by a dot (.), by the numbers 9999, 9998 when the year is missing, and 99 or 98 when the month is missing. When the year of event is missing, the event is assumed not to have occurred, even when the respondent indicated that the event did occur. In other words, the value of year of occurrence is not imputed. For instance one respondent (ID 1569) indicated that she had five children but reported the year (and month) of birth of the youngest child only. Thus it is assumed that she has one child.

Before the month was recorded, the respondent was asked to indicate whether she knows the month of the event. The answer to that question is represented by the variable KAN followed by a number. If the respondent knows the month, KAN* =1, otherwise it is 2. In some rare cases, the respondent did not answer the question and the value of KAN* is missing, represented by 8 or 9. In that case, it is assumed that the respondent does not remember the month. If KAN* is equal to 2, 8 or 9, the month of the event is missing. In some cases, KAN* is equal to one, but the month of event is unknown (value "99"). When the month of event is missing, we imputed the month by assuming that the event occurred in the middle of the year, i.e. in June (month 6). The value of KAN* is not used to determine whether a month is missing. When values are imputed, a flag variable is designated to indicate that the month of occurrence is imputed. Different types of flag variables are used in this report. Flags that indicate a CMC with the month of occurrence imputed starts with FE_, with F denoting flag and E estimation.

The estimation of CMCs raises a number of issues associated with lack of data, in particular data on the month of occurrence of the event. In the remainder of this section, we review the events considered in the OG98.

The OG98 records the number of marriages (AANTKEER) and includes information on at most three marriages: the current marriage and, if applicable, the first and second marriage. The year and month of marriage is recorded and the year and month of marriage dissolution. The survey does not provide direct information on third marriages, but only on the current marriage and at most two previous marriages. It is theoretically possible that the current marriage is not the third but the fourth marriage since no information is provided on the rank of the current marriage. No woman reported more than 3 marriages, however (AANTKEER). If a marriage is dissolved, the reason for marriage dissolution is asked. The reason for the dissolution of the first marriage is denoted by VRG6; it is 1 for divorce and 2

for widowhood. Ten respondents reported the reason for marriage dissolution but not the year (VRG6 known but JRTAL5 missing).
The OG98 includes several questions on cohabitation. A distinction is made between cohabitations that were followed by a marriage and cohabitations that were not followed by a marriage. Ever married women are asked whether they did cohabit *with their future husband* before marriage. Information is collected on cohabitation before current marriage, and previous marriages (first and second marriage, if applicable). All women, including never married women are asked about their first cohabitation and subsequent cohabitations. A woman may report a cohabitation before her current marriage (JRTAL) or her first marriage (JRTAL3) and a first cohabitation (JRTAL10). In that case, the first cohabitation refers to the cohabitation before the cohabitation that preceded the first marriage. That approach to the recording of subsequent cohabitations is confusing and required extensive checking to determine the sequence correctly.

A question is included on the presence of a cohabitation contract. In the original data, the variable CONTRACT is 1 if the woman has a cohabitation contract. Otherwise the value is 2. The year and month of the cohabitation contract is recorded (TYDSTJ and TYDSTM). A total of 333 women indicated that they have a contract while 455 indicated the contrary that they do not. Of those who had a contract, 331 recalled the year of the contract and 291 the month as well. In 40 cases the month was imputed (June).

The OG98 makes a distinction between biological children and adopted children, foster or stepchildren. Women are asked whether they ever had children born alive (OOITLEV) and the number of children (AANTLEV). AANTLEV is coded 1, 2 3, 4 and 5 or more. Fifty-seven women report at least five children. The year and month of birth of biological children is recorded for children born alive, up to a maximum of five children. The fifth child is the youngest child. The year and month of birth of adopted/foster/step children is recorded up to a maximum of three. The sex of the child is recorded for biological children only. In total, 3401 women indicated that they had at least one child born alive. A few women did not recall the year of birth of at least one of the children (see Section 5).

Women were also asked whether they had an adopted child, a foster child or a stepchild (HFTSTIEF). A total of 192 women reported that they had at least one adopted child, foster or stepchild. For each child, the relationship

(adopted, foster or stepchild) is denoted by the variable AARD*, where * is the order of the child (AARD, AARD2, AARD3).

Table 2 shows the data availability for each event. It gives the number of respondents that indicate the year of the event and the month of the event. The table also shows the number of women who indicated that they recall the month of the event. The codes '9998' and '9999' indicate that women do not recall the timing of events. The number is included in the total count of events. The estimation of the timing of the events in CMC required the imputation of the month of occurrence when the information was missing. Table 3 shows for each event, the number of CMCs observed (both year and month are known) and the CMCs imputed (year is known but the month is not). Imputation was required more often for cohabitation than for marriage.

Table 2. Number of respondents reporting year and month of event, various events

	Year (JRTAL*)		Month (MND*)				Recall month (KAN*)			
	Total	9998	9999	Total	98	99	Total	Yes	No	8/9
Birth of respondent	5450	0	0	5450	0	0				
Cohabitation before current marriage	1441	1	1	1292	0	0	1439	1292	141	6
Current marriage	3245	0	0	3245	0	3				
Cohabitation before first marriage	274	0	5	172	0	0	269	172	92	5
First marriage	644	2	8	634	2	22				
End first marriage	644	2	10	632	3	67				
Cohabitation before second marriage	45	0	3	30	0	0	42	30	10	2
Second marriage	60	2	5	53	0	1				
End second marriage	60	3	1	56	1	8				
Husband left	33	0	0	31	0	0				
Current cohabitation	791	0	0	768	0	0	791	768	22	1
First cohabitation	582	0	3	395	0	3	579	395	173	11
End first cohabitation	582	0	4	448	0	3	565	435	126	4
Second cohabitation	72	0	0	48	0	0	72	48	24	0
End 2nd cohabitation	72	0	0	55	0	0	70	53	17	0

Third cohabitation	6	0	0	3	0	0	6	3	3	0
End third cohabitation	6	0	0	3	0	0	6	3	3	0

Table 2, continued

	Year (JRTAL*)			Month (MND*)			Recall month (KAN*)			
	Total	9998	9999	Total	98	99	Total	Yes	No	8/9
Leaving home	2261	0	4	1998	1	5	2257	1998	250	9
Birth second child	2642	0	4	2642	0	4				
Birth third child	884	1	1	884	0	2				
Birth fourth child	221	0	1	221	0	1				
Birth youngest child	57	0	0	57	0	0				

Table 3. Information on occurrence and timing of events based on the original variables

Variable name	*Observed*	*Imputed*	*Total*
Cmcb_op	5450	0	5450
Cmcocm	1292	147	1439
Cmccma	3242	3	3245
Cmcoma1	172	97	269
Cmc1ma	610	24	634
Cmcma1en	562	70	632
Cmcoma2	30	12	42
Cmc2ma	52	1	53
Cmcma2en	47	9	56
Cmchleft	33	2	33
Cmccurco	768	23	791
Cmc1co	392	187	579
Cmco1en	445	133	578
Cmc2co	48	24	72
Cmco2en	55	17	72
Cmc3co	3	3	6
Cmco3en	3	3	6
Cmclh	1992	265	2257
Cmc_k1	3396	0	3396
Cmc_k2	2638	0	2638
Cmc_k3	882	0	882
Cmc_k4	220	0	220
Cmc_k5	57	0	57
Cmc_s1	186	5	192
Cmc_s2	109	3	112

5. SEQUENCE OF TRANSITIONS

5.1 Introduction

Once the state space is specified, the transitions are identified, and the timing of the transitions is expressed in CMC, the individual life path is fully documented. For each individual respondent, the sequence of events in the observation window and the episodes of interest can then easily be traced. That sequence is the basis for the study of pathways leading to an event of interest, e.g. first childbirth. A close look at the sequences reveals a number of measurement problems that result in a few seemingly inconsistent sequences. For instance, not all married women live with their husband. Some cohabit with their partner. Some women start the second marriage before dissolution of the first marriage. The reasons may be diverse and lead to different findings which attributed to the imputation. The problems arise when the sequence of events is not consistent with the timing of the events. This section describes a few issues. The nature of the issues is illustrated in the following examples of individual event histories. Consider respondent 19. The respondent was born in CMC 647, which is November 1953. We consider two life histories, one simple and the other more complex. The simple life history is the marital history given in Box 1. The table is produced by SURVEYLIFE.

Box 1. Marital career of respondent 19

Record :19 CMC of birth : 647
Onset of observation (CMC) : 647
Censoring (CMC) : 1180
Total number of episodes : 6

NOE	TS	Age	Age	State	DURATION	OR	DED
1	647	0	0	NevMar	187	0	1
2	834	187	15	Div1	17	1	2
3	851	204	17	Div2	18	2	3
4	869	222	18	Mar2	121	3	5
5	990	343	28	Div2	120	5	6
6	1110	463	38	Mar3	70	6	8
7	1180	533	44	Censored		8	9

Total duration of observation : 533.00

The first column gives the serial number of the episode (NOE). The first episode of the marital career is spent in the state 'Never Married' (NevMar), which is coded as 1. The episode starts at birth, which is the event-origin. The starting time (TS) is CMC 647. The age at onset is 0 months and 0.00 years. The duration of the episode is 187 months. The episode is terminated at time of marriage (CMC 834) that implies an entry into state 2 (Mar1). States are denoted by a label and a number. The states are:

1. Never married (NevMar)
2. In first marriage (Mar1)
3. Divorced after first marriage (Div1)
4. Widowed after first marriage (Wid1)
5. In second marriage (Mar2)
6. Divorced after second marriage (Div2)
7. Widowed after second marriage (Wid2)
8. In third marriage (Mar3)

In addition to these states, two fictitious states are introduced. They indicate whether a respondent is in the observation state or not. They are introduced for convenience, to describe the starting time and ending time of the observation. The states are denoted by 0 and 9. State 0 is the state 'Not in observation yet' and state 9 is 'No longer in observation (observation terminated)'. When a respondent is not in the observation state, she occupies state 0. Observation starts at exit from state 0. At the end of observation, the respondent enters state 9. The difference between the dates of exit from 0 and entry into 9 is the total duration of observation and the total duration of exposure (observed) to the risk of experiencing events in the marital career. Note that in this illustration, the state space distinguishes first, second and third marriages. It considers the order of the event.

Let us return to the same respondent born in November 1953 (CMC 647). She married at age 15, divorced at age 17, remarried at age 18 and divorced 10 years later. She married a third time at age 38 and remained married until the survey in 1998. At the time of the survey, she was 44 years. The third marriage is the current marriage, i.e. the marriage at interview. The woman was interviewed in April 1998 (CMC 1180). Note that the month of interview is not recorded but imputed. Given the age of the woman at survey time, the survey could have been in any of the months from February to May 1998. The

total duration of observation is 533 months, assuming that observation starts at birth and ends in the estimated survey month.

Now, we consider respondent 19. Suppose that she had a complex partnership or union career (Box 2). When she married at age 15, she lived with her parents. She left home 3 months after marriage (CMC 837), still at the age of 15. The reason for leaving home was coded 'other than marriage or cohabitation'. Strictly speaking that may be correct since she did not leave at the time of marriage, but a few months later. One explanation could be that she reported the date of her civil marriage but that the church wedding took place after the civil marriage. It was not uncommon to have the church wedding after the civil marriage and to remain in the parental home until the marriage was sanctioned by the church. The coding illustrates however the problems that may arise in interpreting the reason for leaving home. We (SPSS syntax) changed the coding from 'other than marriage or cohabitation' to 'marriage' (LEAVEWHY =2). The first marriage was dissolved in CMC 851. In this state space, no distinction is made between divorce and widowhood, although the information is available from the OG98. After the marriage dissolution, she lived alone for some time (DW_alon). A little over half a year after the dissolution of the first marriage, the respondent starts cohabiting (CMC 858) before her second marriage in CMC 869. Before the second marriage was formally dissolved (in CMC 990), she started to cohabit (at CMC 978) and that first cohabitation lasted until CMC 1086, when she started living alone. Seven months later, she started a second cohabitation (at CMC 1093) that the OG98 recorded as cohabitation before the current (i.e. third) marriage. In June 1992 she married for the third time. The woman had no biological children (AANTLEV = 0) and indicated that she does not expect to have any children in the future (VERWMOED =3 and AANTKIND = 0). She is childless by choice (VRIJWIL = 1). At the age of 25, she tried to get pregnant (PROBEER =1 and LFTPROB = 25). She tried for two months (MNDPROB = 2) and did not consult a doctor (UITBLYF = 2). She has two stepchildren (HFTSTIEF = 1; AARD =1; AARD2 = 1). The first stepchild was born in December 1977 (CMC 936), i.e. during her second marriage. She was 24 at that time. The second was born in December 1982 (CMC 996). No information is available on when the stepchildren entered her life. They may be children of the third husband. At the time of survey, none of the two children lived with the respondent (WONEN6 = 2; WONEN7 = 2).

The respondent started cohabiting (CMC 978) with another person while she was formally in her second marriage (dissolved in CMC 990). It is likely that

she and her second husband separated before the formal divorce but no information is available (VRG26 missing). Since cohabitation and marriage are viewed as mutually exclusive, it is assumed that the marriage is dissolved when the respondent started cohabiting with another person, i.e. in CMC 978. Therefore the CMC is adjusted in Box 2. In Box 1 the adjustment was not required since the state space did not specify cohabitation.

The survey also revealed that the woman had a bad relationship with her mother and that the relationship with her father was neither good nor bad (RELVADER = 2; RELMOEDE = 3). Her parents were married and they did not divorce (OOITTROU = 1; OOITSCH = 2). As a child, she experienced family life as neither good nor bad (OUDERVAA = 2). She also indicated that her last relationship ended because of communication problems (RUITELKA). The survey reveals that the woman started her first job in 1970 (not shown in Box 2), which was the year in which the first marriage was dissolved. Currently, she works 70 hours per week (TOTURWK).

Box 2. Union career of respondent 19

Record :19 CMC of birth : 647
Onset of observation (CMC) : 647
Censoring (CMC) : 1180
Total number of episodes (states occupied) : 9

NOE	TS	Age	Age	State	Duration	OR	DE
1	647	0	0	AtHome	187	0	1
2	834	187	15	Married	17	1	4
3	851	204	17	DW_Alone	7	4	5
4	858	211	17	Cohabit	11	5	3
5	869	222	18	Married	109	3	4
6	978[(1)]	331	27	DW_Cohab	108	4	6
7	1086	439	36	Alone	7	6	2
8	1093	446	37	Cohabit	17	2	3
9	1110	463	38	Married	70	3	4
10	1180	533	44	Censored	0	4	7

[(1)] The respondent started cohabitation in CMC 978 and divorced in CMC 990 (see Box 1). In this state space the CMC of divorce has been adjusted to 978.

The OG98 gives the opportunity to study life histories of women in the Netherlands and to assess the impact of experiences and expectations on their lives. The data contain several sequences of events that are seemingly inconsistent. In the remainder of this section, the inconsistencies are

investigated in relation to several main events: leaving parental home; cohabitation, marriage and marriage dissolution, and childbirth. Inconsistencies emerge mainly when the sequence of states and events does not make sense.

We distinguish several subsections. Each subsection is associated with an important event or combination of events. Note that each event may be viewed as being part of a career or domain of life and different careers are combined in the investigation of the OG98. Subsection 5.2 focuses on the event 'leaving parental home'; it considers the timing of the event and the living arrangement after leaving home (destination). In general, the state of destination is based on the reason for leaving home. Subsection 5.3 addresses partnerships: marriage and cohabitation. We show how the dates of marriage and cohabitation and the dates of dissolution of marriage and cohabitation are obtained. Subsection 5.4 addresses childbirth. The final subsection (5.5) shows some inconsistencies between the marital status at time of interview reported by the respondent (variable called, "BURGS_OP") and the final marital status that is derived from the episode data. The section also determines the living arrangement at the time of interview.

5.2 Leaving the parental home

The estimation of the CMC at leaving the parental home is conditioned by the questionnaire. The OG98 records the year and month when the respondent left the parental home differently for (1) women who left home because of marriage or cohabitation and (2) women who left home for another reason. The method therefore depends on the reason for leaving home. If the respondent indicated that she never lived independently (OOITZELF = 2 [no]), the month and year of first marriage or first cohabitation is taken to be the month and year of leaving home. The section of the questionnaire on independent living is skipped. If at the time of survey, the respondent is living independently or if she ever lived independently, she is requested to report the year and month when she started living on her own (Section ZELFSTWONEN: 'no longer living with parents, care-taker, or in an institution'). Three thousand, one hundred and eight-seven respondents indicated that they never lived independently (OOITZELF = 2). Of the women remaining, 1853 indicated that they have lived independently before (OOITZELF = 1) and 410 did not answer the question (OOITZELF missing). Of these 2263 (1853 + 410) respondents, 2257 reported the year they left

home (JRTAL16) and 6 did not. Of these 6, they have lived independently before (ID 926, 2199 and 3409). The other 3 gave no answer to that question and OOITZELF is missing (ID 1582, 4236 and 4377).

Of the 1992 women who reported both year and month of leaving home (Table 2), 202 left home in the same month of first marriage or first cohabitation. Of those 202, 105 left home in the month they married, 96 in the month they started their first cohabitation, and 1 (ID 4827) reported both cohabitation and marriage in the month of leaving home. It is assumed that they left home because of marriage or cohabitation and not for other reasons. A total of 46 respondents left home when they first got married or when they first cohabited (that is before adjustment of the CMC of marriage and/or cohabitation); 19 after a marriage, 25 after cohabitation, and 2 after a marriage and a cohabitation (ID 144 and ID 999). The number of respondents that did not report the CMC of leaving home and did not report a marriage or cohabitation is 440. These respondents were living at the parental home at the time of survey. Two respondents left home after the interview (ID 2240 and ID 3269). Both were single at time of interview. That is not realistic and is a consequence of the estimation of the CMC at interview. To account for the anomaly, the CMC at interview is adjusted to be equal to the CMC of leaving home (see section 6). To easily identify the different cases, the flag variable FTEST_LH is designated. Table 4 shows the values. Note that the number of respondents that report the CMC of leaving home, but did not leave home for marriage or cohabitation is 2002.

The reason for leaving is not recorded in the survey. We created a variable LEAVEWHY to indicate the reason for leaving home. The following reasons are distinguished:

1. Cohabitation
2. Marriage
3. Other reason

Respondents living at the parental home at time of interview are censored and are coded 4 (censored). These reasons are shown in Table 4.

Note that these reasons shown in the present report differ from the reasons for leaving home reported by De Graaf and Steenhof (1999, pp. 22-23). De Graaf and Steenhof consider the reasons given by respondents who indicate that they left home at the right moment or indicated that they would have preferred to leave later or earlier. Of the total number of women who left home at the time of interview (5010 = 5450 - 440), 2680 answered the question on preferred

timing of leaving home (EERDLAT) and missing values (8 or 9) were attributed to 12 women.

Table 4. Flag variable FTEST_LH identifying cases of leaving home

		Count	Reason for leaving
1	CMC leaving home = CMC 1st marriage	105	2
2	CMC leaving home = CMC 1st cohabitation	103	1
3	CMC leaving home = CMC 1st cohabitation = CMC 1st marriage	1	2
4	CMC leaving home > CMC 1st marriage	19	2
5	CMC leaving home > CMC 1st cohabitation	25	1
6	CMC leaving home > CMC 1st marriage > CMC 1st cohabitation	2	2
7	CMC leaving home < CMC first union	1538	3
8	CMC leaving home > CMC interview	2	3
9	No CMC LH, no 1st marriage, no 1st cohabitation	440	4
10	No CMC LH, no 1st marriage, yes 1st cohabitation	396	1
11	No CMC LH, yes 1st marriage, no 1st cohabitation	1441	2
12	No CMC leaving home; CMC 1st marriage < CMC 1st cohabitation	139	2
13	No CMC leaving home; CMC 1st marriage > CMC 1st cohabitation	769	1
14	No CMC leaving home; CMC 1st marriage = CMC 1st cohabitation	8	2
15	CMC leaving home given and none of the above	462	3
	Total	5450	

In 46 cases (4, 5 and 6 in Table 4), the respondent left home after first union. The cases are shown in Table 5. The cases are identified by FTEST_LH = 4 if the first union is a marriage and FTEST_LH = 5 if the first union is a cohabitation. Two cases demand particular attention (ID 144 and 999). Both left home after first cohabitation and a first marriage (FTEST_LH = 6). Respondent 144 reported that she started her first cohabitation in CMC 858, married in CMC 884, and left the parental home in CMC 1054. Respondent 999 indicated that she started cohabitation in CMC 750, married in CMC 787 and left home in CMC 858. It is assumed that the persons left home at time of marriage and because of marriage. The CMC of leaving home was adjusted to their CMC of cohabitation (CMC 858 for ID 144 and 750 for ID 999) and the reason for leaving home is marriage. This adjustment will have an effect on

the inferred living arrangement at the time of survey (see below). Table 5 shows the 46 cases.

Table 5. Respondents who marry or start cohabitation before leaving home

	ID	Flag	Reason for leaving parental home	CMC leaving home	CMC 1st union
1	19	4	1st marriage	837	834
2	43	4	1st marriage	902	845
3	144	6	1st marriage	1054	858
4	232	5	1st cohabitation	1169	1143
5	330	5	1st cohabitation	1028	1014
6	438	5	1st cohabitation	1009	1001
7	450	4	1st marriage	872	849
8	464	5	1st cohabitation	1098	1093
9	653	5	1st cohabitation	1020	1006
10	681	5	1st cohabitation	1062	1050
11	958	4	1st marriage	845	833
12	999	6	1st marriage	858	750
13	1203	5	1st cohabitation	1066	1062
14	1324	4	1st marriage	1038	966
15	1460	4	1st marriage	890	888
16	1539	5	1st cohabitation	990	978
17	1593	5	1st cohabitation	1055	1031
18	1647	5	1st cohabitation	930	894
19	1662	4	1st marriage	960	870
20	1778	4	1st marriage	1018	1009
21	2206	5	1st cohabitation	1050	1040
22	2251	5	1st cohabitation	966	942
23	2344	5	1st cohabitation	1050	1014
24	2346	5	1st cohabitation	973	966
25	2496	4	1st marriage	981	958
26	2504	4	1st marriage	1044	816
27	2585	4	1st marriage	1008	1005
28	2587	5	1st cohabitation	1078	1077
29	2844	5	1st cohabitation	1078	1049
30	2961	5	1st cohabitation	1038	963
31	3372	5	1st cohabitation	1140	1101
32	3379	4	1st marriage	855	852
33	3450	4	1st marriage	1120	1116
34	3659	4	1st marriage	867	838

Table 5, continued

	ID	Flag	Reason for leaving parental home	CMC leaving home	CMC 1st union
35	3867	4	1st marriage	906	900
36	3919	5	1st cohabitation	1092	1091
37	3924	5	1st cohabitation	1040	988
38	4341	5	1st cohabitation	1158	1157
39	4455	4	1st marriage	990	989
40	4553	5	1st cohabitation	1071	1070
41	4646	5	1st cohabitation	976	736
42	4789	4	1st marriage	819	810
43	4900	5	1st cohabitation	1134	1130
44	4971	4	1st marriage	908	906
45	5029	5	1st cohabitation	1122	1115
46	5249	4	1st marriage	1020	996

ID 144 and 999 started cohabitation and married before leaving parental home

To sum, we summarize four steps involved in transforming CMCLH to CMCLEAVE. In addition, we consider special cases. The date of leaving home in CMC (CMCLEAVE) is obtained as follows:

i. If a respondent left home for reasons other than marriage or cohabitation, the year and month of leaving home is given (JRTAL16 and MND16). The CMC is easily calculated. If MND16 is missing, it is imputed to be 6 (June).

ii. If the respondent left home because of cohabitation or marriage, the year of leaving home (JRTAL16) is missing. The year (and month) of leaving home is equal to the year (and month) of first marriage or first cohabitation, whichever came first.

iii. If the JRTAL16 is not missing and the year of first cohabitation or first marriage is before the year of leaving home, the respondent lived at the parental home at time of union formation. In that case CMC of first union experience (CMFUNION) is taken as CMC at leaving parental home (CMCLEAVE). A total of 46 respondents were in that situation; 25 started cohabitation while remaining at home, 19 married before they left home, and 2 started cohabitation and married before they left the parental home.

iv. If JRTAL16 is missing and the respondent is not married or cohabiting, she is living at the parental home at the time of survey. A total of 440 respondents lived at the parental home at time of interview.

The following special cases are considered.

i. One woman (ID 5435), who was 53 years old when interviewed, reported three marriages and that she was divorced at time of interview (BURGS_OP). She did not report when she left the parental home (JRTAL16 missing) and did not report any date of marriage or marriage dissolution; she reported the years and months of birth of her two children, however. In the adjusted data set, it is assumed that she remained at her parental home, although one could equally assume that she lives independently. In that case, the year of onset of independent living remains unknown.

ii. Respondent 3698 indicated that she has lived independently before, but left home in 1987 and started cohabitation in 1987. The month of leaving home and the month of cohabitation are not known. It is assumed that both events took place in June 1987.

iii. Respondent 4646 misreported the year in which she started cohabitation. The year reported is 1961 and the month is April. However, the respondent was born in September 1961 and she indicated that she left home in April 1981 (JRTAL16 and MND16) and married (current marriage) in September 1981 (JRTAL2 and MND2). It is therefore reasonable to assume that she started cohabitation before the current marriage in April 1981. The CMCOMA1 has been changed to 976 (in the syntax module for estimation of CMC of leaving home) and a flag variable F_BEWARE has been added.

Table 6 shows the reasons for leaving home.

Table 6. Reasons for leaving parental home

	Frequency	Percentage
Cohabitation	1295	23.7
Marriage	1713	31.5
Other	2002	36.7
Staying at parental home (censored)	440	8.1
Total	5450	100.0

5.3 Union formation and dissolution

The OG98 gives information on up to three marriages, three cohabitations before a marriage, and four cohabitations that do not lead to marriage. The majority of women have 1 or 2 partnerships, in the form of either marriage or

cohabitation. Table 7 shows the number of partnerships. The figures are based on the Public Use Data File BOAV98.SAV. AANTKEER denotes the number of marriages. NOGMEERx indicates whether the respondent has more than x cohabitations. For instance, NOGMEER3 is a 0-1 variable, which is 1 if the respondent had more than 3 cohabitations (excluding cohabitations that lead to marriage and with the future husband). This section deals with three events: marriage, cohabition, and end of marriage or cohabitation.

Table 7. Number of partnerships reported in OG98

Number of Partnerships	All	Marriages	Cohabitations	AANTKEER*	NOGMEER*
0	904	1774	2836	n.a.	n.a.
1	4546	3676	2614	n.a.	72
2	1928	244	513	230	6
3	482	12	66	24	2
4	137	n.a.	4	n.a.	n.a.
5	28	n.a.	1	n.a.	n.a.
6	8	n.a.	n.a.	n.a.	n.a.
7	1	n.a.	n.a.	n.a.	n.a.

Variables are taken from:
AANTKEER = "how many times have you married?"
NOGMEER (2/3) = "any other cohabitations after second/third cohabitation?"

a. Marriage

In OG98 a distinction is made between current marriage and previous marriages. When the current marriage is the first marriage, information is given for the current marriage and no separate information on the timing is given for the first marriage. In other words, the year of first marriage (JRTAL4) is given only when the current marriage is not the first. If the current marriage is the first, the year of current marriage (JRTAL2) is the year of first marriage. The number of marriages reported by the respondent is AANKEER. The number of marriages may also be estimated from JRTAL2, JRTAL4 and JRTAL7. The estimated number of marriages is denoted by NMARRIAG. It is the number of marriages for which the year for marriage is known. If JRTAL7 is less than JRTAL2, then the current marriage is assumed to be the third marriage. It could be a higher-order marriage since the year of marriage is requested for the current marriage and the first two marriages only. No woman reported more than 3 marriages, however.

Several respondents who married only once did not report the year of current marriage (JRTAL2) but the year of first marriage (JRTAL4). The OG98 records marriage dissolution, if applicable, for the first and second marriages only. The year and month of marriage dissolution is JRTAL5 and MND5 in case of the first marriage and JRTAL8 and MND8 in case of the second marriage. Women who reported a marriage dissolution were asked how many months they were separated from their husband before the formal dissolution of the marriage (VRG25 in case of dissolution of first marriage and VRG26 in case of dissolution of second marriage). In OG98, 644 women reported the end of the first marriage (JRTAL5 not missing). It includes 12 women with JRTAL equal to 9998 or 9999. For 632 women, the year in which the first marriage ended is known. In that group, 542 reported the number of months the couple was separated before the formal dissolution. The codes 998 and 999 are treated as missing (36 cases). The reason for dissolution was a divorce in all cases that include information on the duration of separation and the reason for dissolution (213 cases; cross-classification of VRG25 by VRG6). Most women left the question on reason for marriage dissolution unanswered. We may assume that divorce was the reason for the dissolution of the marriage. Of the 632 women with known CMC at first marriage dissolution, 539 reported the duration of separation before the formal dissolution of the marriage (cross-classification JRTAL5 by VRG25). On the basis of the information on duration of separation before divorce, one may estimate the CMC of separation.

Of the 60 women who reported a dissolution of the second marriage (JRTAL8 not missing, but may be equal to 9998 or 9999), 45 reported the number of months the couple was separated before the formal dissolution (cross-tabulation JRTAL8 by VRG26). Missing information is coded 998 and 999 (8 cases). Of the 45 women, only 15 indicated that the marriage dissolution was because of divorce (cross-classification of VRG26 by VRG10). From these variables, the CMC of separation before the divorce that ended the second marriage has been estimated.

Table 8 shows the number of respondents by marital status at time of survey, as reported in the OG98 (BURG_OP). A total of 3680 women married at least once before the survey and 3245 women were married at time of survey (currently married). All reported their year of marriage. As a consequence, that figure is the same whether the number of married women is produced by BURG_OP or by JRTAL2, JRTAL4 and JRTAL7. Of the currently married

women, 3036 were married for the first time, 194 were in their second marriage and 15 in their third marriage.

Table 8. Marital status at time of survey

	Frequency	Percentage
Married	3245	59.5
Divorced	378	6.9
Widowed	57	1.0
Never married	1770	32.5
Total	5450	100.0

Of the 378 women who were divorced at time of interview (BURG_OP) and the 57 widows (BURG_OP), 390 reported to have been married once and 36 have been married twice and 9 three times (Table 9). Table 9 is a cross-tabulation of AANTKEER and BURG_OP.

Table 9. Number of marriages by marital status at time of survey

Number of marriages	Marital status at time of survey				
	Married	Divorced	Widowed	Never Married	Total
0				1770	1770
1	3036	338	52		3426
2	194	32	4		230
3	15	8	1		24
Total	3245	378	57	1770	5450

Four women (3 divorced and 1 widowed at time of survey) reported their number of marriages (AANTKEER) but did not report the year of first marriage (JRTAL4 missing). They are ID 363 (widowed, married twice), 929 (divorced, married once), 1545 (divorced, married once) and 5435 (divorced, married three times). The CMC of first marriage is estimated for 3676 ever married women.

At least one woman (ID 1587) reported more than one marriage but indicated only the year of the current marriage. She reported two marriages (AANTKEER=2). She also reported the year of current marriage (JRTAL2) but did not recall the year of first marriage (JRTAL4 = 9998). She did not

report the marriage dissolution either. It is assumed that she married only once and is currently married.

Of the respondents, 24 women reported three marriages (AANTKEER=3). Not all these women gave complete information on the timing of marriage and marriage dissolution. The year of the third marriage is not recorded by the OG98, unless the third marriage is the current marriage. Twelve women reported the year of third marriage; which was their current marriage. For 8 women the third marriage was dissolved at time of survey; 7 were divorced (ID 43, 1449, 1950, 2677, 3570, 4437, 4530, and 5435) and 1 was widowed (ID 5134). They reported the years of the first and second marriages as requested by the survey. The year of third marriage was not requested and remains unknown. Four women provided incomplete information. One woman (ID 5435; divorced at survey) with three marriages did not report any date of marriage or marriage dissolution. She did report the date of birth of two children, however.

Another woman (ID 2874; married at survey) gave only the year of current marriage and did not report the dissolution of the first marriage, the second marriage, and the dissolution of the second marriage. We assume that respondent 2874 was married only once, although she reported three marriages (AANTKEER =3). Note that we did not change the value of AANTKEER, but introduced a new variable NMARRIAG to denote the number of marriages for which the year of marriage is available. A comparison of AANTKEER and NMARRIAG reveals therefore the differences and the cases responsible for the differences. A third woman (ID 1570; married at survey) reported the year of first marriage and the current marriage (third). She did not report the year of dissolution of first marriage (JRTAL5), the year of second marriage (JRTAL7), and the year of dissolution of the second marriage (JRTAL8). Neither did she indicate when the marriages were dissolved. Her case should not be used for marital status analysis. Finally, respondent 1575 (married at survey) reported the year of first marriage and the year of first marriage dissolution. She also reported the year or current marriage. However, the respondent did not provide information on the year of the second marriage and the second marriage dissolution. We assume that she married twice, although she indicated three marriages.

Respondent ID 825 married three times and did not cohabit. She reported the same date for the end of the first marriage and the end of the second marriage

(February 1985). The second marriage however started in September 1984 while the current marriage (third marriage) started in March 1985. It is evident that the end of first marriage (JRTAL5 and maybe MND5), which ended in a divorce, is misreported. We assume that the second marriage ended in February 1984. This assumption is entered in the syntax (see 4_CMC events.sps). These cases are indicated by the flag variable F_BEWARE.

Table 10 shows the number of marriages reported (AANTKEER) by the number of marriages for which the date of marriage is given.

Table 10. Number of marriages: reported and number based on year reported

		NMARRIAG (Number of marriages)				
		0	1	2	3	Total
AANTKEER	0	1770				1770
(How many	1	2	3424			3426
times have you	2	1	7	222		230
married?)	3	1	1	10	12	24
	Total	1774	3432	232	12	5450

b. Cohabitation

A distinction is made between cohabitations that were followed by a marriage and cohabitations that were not followed by a marriage (see Table 1). Ever married women were asked whether they cohabited *with their future husband* before marriage. The OG98 assumes that the marriage following cohabitation is with the same partner. No provision is made in the questionnaire for marriage to a different partner. Seven different cohabitations are named:

a. cohabitation before current marriage (Cmcocm);
b. cohabitation before first marriage (Cmcoma1);
c. cohabitation before second marriage (Cmcoma2);
d. current cohabitation (Cmcurco);
e. first cohabitation (Cmc1co and Cmcco1);
f. second cohabitation (Cmc2co Cmcco2); and
g. third cohabitation (Cmc3co and Cmcco3).

The number of cohabitations (NCOHABIT) is determined by the values of JRTAL, JRTAL3, JRTAL6, JRTAL9, JRTAL10, JRTAL12 and JRTAL14. If a value is not missing, then it means that the respondent reported the year of

onset of cohabitation. The number of cohabitations is denoted by NCOHABIT. Of the 5450 women interviewed, 2836 never cohabited, 2101 cohabited once, 447 twice, 62 three times, 3 four times and 1 five times. The respondent who reported to have cohabited 5 times (ID 171) started the third cohabitation in the same month as the cohabitation before the first marriage (JRTAL14 = JRTAL). It is likely that the respondent misunderstood the questions and reported cohabitation irrespective of whether it was cohabitation that led or did not lead to marriage. The data have not been adjusted, however. Table 11 shows the cross-classification of number of cohabitations and number of marriages. Of the 904 respondents who never cohabited and did not marry, 440 were living at home at time of survey (cmclh is missing) (Table 6).

The month in which cohabitation ended is reported as well as the reason for the termination. The following reasons are distinguished: marriage (for cohabitations before current or a previous marriage), separation, death of the partner, or other reason. A total of 582 women who ended the first cohabitation for reasons other than marriage reported the reason for ending the cohabitation (VRG11). Seventy-two women who ended a second cohabitation without marrying the partner reported the reason for ending the cohabitation (VRG12) and 6 women indicated why they ended the third cohabitation (VRG13).

Table 11. Number of cohabitations by number of marriages

		NMARRIAG (Number of marriages)				
		0	1	2	3	Total
NCOHABIT (Number of cohabitations)	0	904	1903	27	2	2836
	1	715	1282	102	2	2101
	2	137	223	84	3	447
	3	18	23	17	4	62
	4			2	1	3
	5		1			1
	Total	1774	3432	232	12	5450

For each respondent, the number of marriages NMARRIAG and the number of cohabitations NCOHABIT were estimated based on the years of marriages and the years of onset of cohabitations that are reported or estimated. The total number of unions is the sum of marriages and cohabitations. The number of these events is shown on Table 12. Eight respondents have 6 or more unions.

Table 12. Number of marriages, cohabitations, and unions

Number	Number of marriages NMARRIAG	Number of cohabitations NCOHABIT	Number of unions NUNIONS
0	1774	2836	904
1	3432	2101	2618
2	232	447	1446
3	12	62	345
4		3	109
5		1	20
6			7
7			1
Total	5450	5450	5450

Table 13. Number of marriages and cohabitations for women with six or more unions

Number	ID	Age at time of survey	Number of marriages NMARRIAG	Number of cohabitations NCOHABIT	Number of unions NUNIONS
1	19	44	3	3	6
2	171	42	1	5	6
3	343	42	2	4	6
4	438	30	2	4	6
5	3641	40	3	4	7
6	4430	49	3	3	6
7	4848	44	3	3	6
8	4995	40	3	3	6
Total	8	8	8	8	8

The individual data are shown in Table 13. Table 12 shows that 3 women had 4 cohabitations (ID 343, 438 and 3641) and 1 woman had 5 cohabitations (ID 171). That finding differs from the number of cohabitations that may be derived from the variable HOEVAAK3 that was included in the original OG98 data set. Two women reported that they had more than 3 cohabitations (NOGMEER3 = 1) (ID 171 and 3840). The first reported 1 additional cohabitation (HOEVAAK3 =1) and the second 2 (HOEVAAK3 = 2). Respondent ID 3840 who did not marry and did not cohabit before a marriage, recorded 3 cohabitations, as requested. A comparison of HOEVAAK3 and NCOHABIT indicates the underreporting of cohabitations that are not

followed by a marriage (HOEVAAK3 refers to cohabitations that are not followed by a marriage).

c. Adjustment of the end of marriage and cohabitation

The OG98 records a maximum of 3 marriages, including the current marriage. The time at the end of the first marriage is given by JRTAL5 and MND5. The reason for ending the first marriage is given by the variable VRG6 with 1 for divorce and 2 for widowhood. The information is the basis for the calculation of the CMC at divorce and widowhood. A similar procedure is followed for the end of the second and third marriage. The reason for ending the second marriage is given by VRG10.

Women may start a new partnership or may start living alone before the partnership is fully terminated. As a consequence, the measurement of the CMC at dissolution of partnerships (marriage or cohabitation) is problematic. Consider a marriage that ends in a divorce. The date of divorce is determined by a legal procedure. Before the legal procedure is over, a couple may separate and the woman may live alone or with another partner. In the section on marriage and marriage dissolution, we noted that, of the 632 women who knew the year in which the first marriage ended, 539 reported the number of months the couple was separated before the formal dissolution (cross classification JRTAL5 and VRG25). Of the 60 women who reported the dissolution of the second marriage, 45 reported the number of months the couple was separated before the formal dissolution. Note that of the 60 women who reported the dissolution of the second marriage, 4 did not report the year of dissolution (year coded as 9999).

Whether an adjustment needs to be made to the CMC of marriage dissolution depends on the treatment of separation. If the state space distinguishes ‘separated’ as a living arrangement, separation is possible. In that case, a divorce that occurs after a woman starts living with another partner is perfectly reasonable. No adjustment is required. If, on the other hand, being married but separated is not a living arrangement, an adjustment is needed. Since in the previous section, separation is not distinguished in the state space, an adjustment is made when a woman reports cohabiting before a formal divorce. Two options exist. The first is to assume that the marriage is dissolved at the time of the divorce. The second is to assume that the marriage ends at the start of cohabitation. In this report (syntax), the second option is selected. It is assumed that the marriage ended at the start of the new union. In

other words, if a woman starts a new union without first ending the previous union, the END of the previous union, expressed in CMC, is recoded to be the same as the BEGINNING of the subsequent union. A similar procedure is used to determine the end of cohabitation. The flag variable FA_CO identifies the cases that called for an adjustment.

Consider respondent with ID 1879. At time of interview, the woman was 39 years old and married. She left the parental home in July 1979 (CMC 955) for reason of marriage. The marriage was dissolved in August 1983 (CMC 1004) and the reason was divorce (VRG6 = 1). However, she started her first cohabitation in January 1983 (CMC 997), i.e. seven months before the divorce. She indicated that she separated 6 months prior to the divorce (VRG25 = 6). Since separation was not a transition in the state space, we changed the CMC at divorce to 997 and attached a flag variable FA_co1 = 1. The cohabitation ended in January 1984 (CMC 1009). About a year and a half later, in August 1985 (CMC 1028), she started a new cohabitation with the partner she married in February 1989 (CMC 1078). At the time of interview she was still living with that partner. The case illustrates the significance of the state space for the study of event histories. A total of three women started cohabiting before the formal dissolution of the first marriage.

In OG98, 644 women reported the end of the first marriage (JRTAL5 not missing). It includes 12 women with JRTAL equal to 9998 or 9999, which indicates a missing value. For 632 women, the year in which the first marriage ended is known. In 42 cases, the date is inconsistent with the other events. For 590 marriages, the time ended is taken from the observed end of first marriage. For 54 first marriages, the ending time is taken from the adjusted end of first marriage (42 cases with inconsistent dates and 12 cases with missing dates). Sixty women indicated the end of a second marriage, including 4 women who did not report the year of the event. Fifth-six women knew the end of the second marriage (cmcma2en). The ending time of the second marriage is recorded for 12 women. In 9 cases, the ending time is equated to the starting time of a new union. The ending times of cohabitation are also recoded if necessary. Unlike the end of a marriage, there are three possibilities for women to end a cohabitation. The first is a discontinuation of the partnership, the second a marriage, and the third the death of the partner.
Some women who reported that they started cohabition before the legal divorce also reported the number of months of separation from their husband before the divorce (VRG25 and VRG26). When women started cohabitation before the date of divorce, the CMC of marriage dissolution was adjusted to

the CMC of cohabitation. This adjustment procedure leads to new problems. Ten women reported that they started cohabiting before they separated from their husband (ID 340, 920, 936, 1493, 1879, 2141, 3273, 4794 and 5074). Of the women who reported that they started cohabiting before the end of the second marriage, one separated from her husband AFTER the onset of cohabitation (ID 1983). Part of the problem could be related to the imputation of the CMC. That does not seem to be the case, however (see values of flag variables associated with estimation of CMC in module 4_CMC events.sps). We adjusted the CMC of separation if the woman reports cohabitation before separation (and divorce). The adjustment is carried out in module 6_Cohabitation sequence.sps. The adjusted cases are identified by the flag variables F_MA1SE for adjustments of the CMC of the separation before the end of first marriage, and F_MA2SE for adjustments of the CMC of the separation before the end of the second marriage. The last variable is not needed because there are no cases, but is included for completeness.

The discussion illustrates the importance of the state space. Recall that the states are mutually exclusive, which means that a person may occupy only one state at a time. In a number of cases, the reported sequence of events cannot be handled by the state space that has been defined. Consider the state space in Figure 3 and assume that some sequences are not allowed. Examples may include marriage before leaving home, cohabitation before leaving home, cohabitation before the dissolution of a marriage, and remarriage without a marriage dissolution. Whether these sequences are possible or not depends on the state space and the transitions that are considered feasible given the state space. When a marriage or cohabitation occurs before leaving home, the CMC at leaving home is equated to the CMC of marriage or cohabitation. In two cases (ID 2874 and 5251), the CMC of leaving home was missing (original 9999) but the respondent did marry (in 1986). It is of course possible that a person marries without leaving the parental home. These two respondents provided incomplete information on other events too (ID 2874 lacks information on JRTAL2, JRTAL5, JRTAL7 and JRTAL8; ID 5251 lacks information on JRTAL2, JRTAL4 and JRTAL5). Since the states in the state space are mutually exclusive, which means that a person may occupy only one state at a time, one may not live at parental home and be married at the same time. The two persons were therefore considered to be ‘married’. The number of respondents living at home at time of interview is 440.

5.4 First birth and higher-order births

Of the 5450 women in the sample, 3401 indicated that they had at least one child born alive (OOITLEV = 1). Not all women reported the date of birth of the children. One woman (ID 1432) with four children did not report the year and month of birth of any of the children. She reported the sex of the children, however. Another woman (ID 3904) did not report the year and month of birth of the first child (GEBJR_K1 = 9999, GEBMD_K1 = 99), but reported the sex, and reported the year and month of birth of the second and third children. One woman (ID 1569), who had no less than five children, reported the year and month of the youngest child only. The same woman did report the age of marriage and the age at which she tried to get pregnant (LFTPROB = 25 and LEEFT2K = 28). Two women (ID 166 and 168) indicated that they had two children but did not report the year and month of birth of the children. As a result, the year of birth is known for 3996 first children.

One woman (ID 4803) had one child born in March 1998 (CMC 1179). The survey month (CMCINT) was estimated at February 1998 (CMC 1178). The survey month is adjusted (see Section 6).

The study of the sequence of births reveals the number of women with twin births. No woman had triplets. The analysis indicates some inconsistencies in the data. For 47 women, the first delivery was a twin birth (CMC_K1 = CMC_K2). For 37 women, the second delivery was a twin birth (CMC_K2 = CMC_K3). The third delivery was a twin birth for 9 women and the fourth for 3 women. One woman (ID 1025) with four children had two twins and one (ID 3105) with five children had two twins (the first delivery and the third delivery; the second delivery was a singleton).

The inconsistencies relate to three women. The first case (ID 1569) has been discussed earlier in this section. The second is ID 3571, who married in June 1968, reported the birth of a first child in July 1974 (CMC 895) and the birth of a second child in April 1974 (CMC 892). The third case (ID 4098) is a woman who reported four children (AANTLEV = 4) and the years of birth of three of the four children. The year of birth of the second child is missing (GEBJR_K2 = 9999). A flag variable is designated to these cases: F_K12345 = 1. They should be omitted in studies of the fertility career of women.

Women also reported adopted children, foster children or stepchildren (HFTSTIEF). A total of 192 women reported that they had at least one

adopted child, foster child or stepchild. In 150 cases the oldest child was a stepchild, in 18 cases it was an adopted child and in 24 cases it was a foster child. Five could not recall the year of birth of the oldest child and one woman reported the year of birth of a foster child as 1872 (ID 2059). The case is indicated by the flag variable F_BEWARE. Since it is likely to be a typing error, the year was changed to 1972 (in Module 4 of the syntax, which estimates the CMCs). The number of women with two adopted, foster or stepchildren is 112 (92 stepchildren, 15 adopted children and 5 foster children). Three did not recall the year of birth of the second child. Twenty-one women had 3 children (17 stepchildren, 1 adopted child and 3 foster children).

5.5 Marital status and living arrangement at time of interview

The final subsection compares the marital status recorded at interview and represented by the variable BURGS_OP, and the marital status derived from the sequence of events and the timing of events (CMCs). The marital status based on the sequence of events is the marital status of the woman after the LAST event prior to the interview. The comparison is carried out AFTER the CMC at interview was adjusted to remove inconsistent cases with an event after the interview date (see next section).

Table 14. Marital status at the time of interview (BURGS_OP) and marital status derived from partnership career

	Based on partnership career		BURGS_OP	
Marital status	Frequency	Percentage	Frequency	Percentage
Married	3247	59.6	3245	59.5
Divorced/ Widowed	431	7.9	378 57	6.9 1.0
Single	1772	32.5	1770	32.5
Total	5450	100.0	5450	100.0

The comparison is shown in Table 14. The syntax uses the imputed interview date and determines the marital status. The syntax is given in the last part of Module 8. We found 4 cases where the marital status reported by the respondent differs from the marital status obtained from the partnership career. The individual cases are shown in Table 15. The four cases that remain

after the interview date is adjusted to accommodate cases with an event after interview date (4 cases; see next section). It is a consequence of the need to estimate the interview date since that date is not included in the Public Use File (for reasons of privacy). Other cases include ID 1582 where the inconsistency pertains to the beginning and end of marriage (CMCs are 978 and 966) and ID 1545, 2856, 5435, where the occurrence of the event is known but the timing is not. These cases make it difficult to determine the marital status at the time of interview from the sequence data.

Table 15. Case summary of inconsistencies between marital status at time of interview reported by respondent and marital status derived from the union career

Case	ID	Age at survey	Marital status at interview based on CMC (INTSM1)	Marital status (BURGS_OP)
1	1545	35	Never married	Divorced
2	1582	40	Married	Widowed
3	2856	48	Married	Widowed
4	5435	53	Never married	Divorced

The living arrangement at the time of survey is not reported directly by the respondent. It is derived from the CMCs at events. The living arrangement is determined by the last event prior to the survey (based on CMC at event). The results are shown in Table 16. Sixteen living arrangements are distinguished. At time of interview, 986 women lived alone. The status just before they started living alone is also considered. Of the 986 women, 464 lived in the parental home, 172 were cohabiting for the first time and 56 cohabited for the second time, 257 were in their first marriage and 37 in their second marriage. The syntax is shown in Module 10 in Appendix A.

Table 16. Living arrangement at time of survey

A. Aggregate categories	Frequency	Percentage
Parental home	440	8.1
Alone	986	18.1
Cohabitation	782	14.3
Married	3242	59.5
Total	5450	100.0

B. Detailed categories	Frequency	Percentage
Alone; left home (1)	464	8.5
Alone; end first cohabitation (2)	172	3.2
Alone; end 2nd cohabitation (3)	56	1.0
Alone; end 1st marriage (5)	257	4.7
Alone; end 2nd marriage (6)	37	.7
First cohabitation (8)	609	11.2
Second cohabitation (9)	161	3.0
Third cohabitation (10)	6	.1
Fourth cohabitation (11)	1	.0
Fifth cohabitation (12)	5	.1
First marriage (13)	3038	55.7
Second marriage (14)	192	3.5
Third marriage (15)	12	.2
Living at parental home (16)	440	8.1
Total	5450	100.0

6. MONTH OF INTERVIEW

The OG98 was carried out between February and May 1998. The month of interview was omitted from the Public Use File. Although the information is known to be essential to determine the duration of exposure to the risk of demographic events and to process censored cases correctly, other renowned statistical agencies adopt the same practice and omit the month of interview (see e.g. Mills, 2000, p. 260 discussing data released by Statistics Canada). The OG98 includes the age of the respondent at interview, however (age in completed year). We imputed the month as part of the conversion of the Public Use File into an event history data structure. The method was suggested by drs. José Dias as part of an assignment in the course on 'Multistate demography'. The month of interview is denoted by "CMCINT_O" (century month code).

At the outset, it is assumed that each month of the period from February to May is equally likely to be the month of interview, provided that the age of the respondent at that month coincided with the reported age. The first step of the imputation procedure determines the age of the respondent at each month of the period of interview. The age is the difference between the CMC of the month of interview and the CMC of the month of birth, divided by 12. When the estimated age differs from the reported age, that month could not have been the month of interview. That month is inconsistent with the reported age.

Table 17. Sequence of consistent and inconsistent survey months

FLAG_TOT	Frequency	Percentage
0000	139	2.6
1000	26	.5
1100	154	2.8
1110	290	5.3
0001	196	3.6
0011	316	5.8
0111	396	7.3
1111	3933	72.2
Total	5450	100.0

A month that is inconsistent is coded 0 and a month that is consistent with the reported age is coded 1. To each month a flag of 0 or 1 (see syntax) was designated. Dias defined a single variable to indicate the sequence of

consistent and inconsistent months. The variable is FLAG_TOT. Table 17 shows the sequences.

Each sequence xxxx corresponds to the sequence February, March, April and May. Here, 0 indicates that the month is inconsistent with the reported age. For a total of 3933 women any month can be chosen to be the month of interview (1111). For 26 women the information only remains consistent assuming that the survey was in February [MOTH_SRV = 2 (February)]. Other months are not possible given the reported age. For 196 women, the survey month is consistent with the reported age only if the survey was in May 1998 [MOTH_SRV = 5 (May)]. In these two cases, the flag variable associated with the survey month (FLAG_SRV) indicates that the survey month is *exact*.

Table 18. Sequence of feasible survey months by month of birth

	Month of birth												
FLAG_TOT	Jan	Febr	March	April	May	June	July	Aug	Sept	Oct	Nov	Dec	Total
0000					106	33							139
1000		26											26
1100			154										154
1110				290									290
0001				196									196
0011			316										316
0111		396											396
1111	449				369	448	451	435	463	441	437	440	3933
Total	449	422	470	486	475	481	451	435	463	441	437	440	5450

If two months are feasible, the survey month is sampled from a binomial distribution (the cases 1100 and 0011). If more than two months are feasible, the survey month is sampled from a multinomial distribution (1, pi). Since the multinomial distribution is not available in SPSS, Dias used the quantiles of a uniform distribution (e.g. in the case of Bernoulli, it would be < 0.5 and => 0.5; pi is equal to one divided by the number of feasible months). The flag variable indicates that the survey month is *imputed*. Table 18 shows the sequence of feasible survey months by month of birth of the woman. When the data are inconsistent (0000), June/July is assumed to be the survey month. The flag variable indicates *inconsistency*. Dias estimated that 106 women were interviewed in June and 33 in July, unless the age is misreported and the birthday is after the interview. Note that, although Statistics Netherlands

reported that the survey was held during the months February to May, 139 either misreport their age at survey or the survey month was outside of that range. Table 19 shows the survey months. Note that a new application of the procedure (syntax) does not reproduce Table 19 exactly because a random number is generated to select a month from a range of feasible months. Every application of the syntax with the random number generator leads to slightly different results. When multiple months are feasible, the procedure assumes that each of the feasible survey months is equally likely. The SPSS syntax to generate estimates of the survey month is presented in the Appendix (3_Surveymonth.sps). The estimated month at interview imputed by Dias is the variable MOTH_SRV. The CMC at time of interview is CMCINT_O = 98*12+MOTH_SRV. It is stored in the file 2_og98_cmcint.sav with only column: the estimated month of interview (CMCINT_O). Syntax to merge that file with the larger data file is included (31_Match files INT.sps).

Table 19. Estimated survey months

Survey month	CMC	Frequency	Percentage
February	1178	1177	21.6
March	1179	1321	24.2
April	1180	1364	25.0
May	1181	1449	26.6
June	1182	106	1.9
July	1183	33	.6
Total		5450	100.0

The survey month determined by an imputation procedure based on a single variable (age) may be inconsistent with the CMCs that are estimated for the different events. Some events may occur after the estimated survey month. The inconsistencies are determined after the estimation of the CMCs and the adjustments of the CMC of marriages and cohabitations. When an event is identified to have occurred after the estimated CMC at time of interview, the CMC at interview (CMCINT_O) is adjusted. Eight cases were identified (Table 20). One woman (ID 4803) had her first child after the estimated survey date. That date is adjusted. In addition, one CMC (ID 1589) is adjusted to a month outside of the 'plausible' range from February to July 1998. In that particular case, the woman reported the dissolution of her first marriage in September 1998 (see JRTAL5 and MND5). It is probably a recording error since that particular woman started her current cohabitation in November 1997. The CMC at the end of the first marriage is adjusted to be the CMC at

the onset of the current cohabitation. Consequently, the CMC at time of interview did not need to be adjusted (CMCINT_O = 1178). Therefore that case is not shown in Table 20.

Table 20. Inconsistent CMC at interview

	ID	Initial CMC CMCINT_O	Adjusted CMC CMCINT
1	869	1179	1180
2	1386	1178	1179
3	1464	1178	1179
4	2240	1178	1179
5	2377	1178	1181
6	3269	1178	1182
7	4928	1178	1180
8	4803	1178	1179

7. EVENT HISTORY DATA FILE

The event history data file is derived from the original Public Use File according to the following steps:

1. Translate the variable labels and value labels from Dutch into English.
2. Select a subset of variables for event history analysis.
3. Impute the survey month.
4. Estimate the CMC of the relevant events and determine the number of marriages and cohabitations.
5. Determine the CMC of first, second and third marriages.
6. Determine the CMC of first and subsequent cohabitations.
7. Estimate the CMC of leaving parental home.
8. Remove inconsistencies in CMCs of marriages, cohabitations, and leaving home.
9. Determine singleton, twin and triplet births.
10. Determine the living arrangement at the time of survey.
11. Save the results in an event history data file.

Each step has a separate SPSS syntax. The syntax is organized in 11 modules, with one module having two variants. Each module implements a particular task and builds on the previous module. The input data to the sequence of modules consists of the original Public Use File (BOAV98.SAV) that is distributed by the Scientific Statistical Agency (WSA) of the Netherlands Organization for Scientific Research (NWO). The sequence of modules is shown in Box 3. The SPSS syntax for the marital status analysis and the syntax for the union status analysis are shown in the Appendix.

The output of the sequence of modules is a data file that contains the dates of the events in CMC, duly checked for inconsistencies in sequence and timing of the events. The file is named NLOG98_F_CMC.SAV. In addition to the CMC values, the file contains sets of variables that indicate how the CMCs were obtained. They are flag variables. They are easily identified in the syntax since they start with F_ (Flag variable), FE_ (flag variable related to the estimation of CMCs), or FA_ (Flag variable in the module that Adjusts inconsistencies).

The CMCs and the flag variables that are created, and additional variables that are considered relevant and included in the event history data file are listed

below. The list of variable names may also be obtained by running a simple SPSS syntax that includes one line: display labels.

a. General variables newly created or copied from original file

-	ID	Identification number [*new*]
-	WFT	Sample weight
-	LFT_OP	Age of respondent at survey
-	BURGS_OP	Marital status at survey
-	NMARRIAG	Number of marriages [*new*]
-	NCOHABIT	Number of cohabitations [*new*]
-	NUNIONS	Number of unions [*new*]
-	LEAVEWHY	Reason for leaving home [*new*]
-	AANTLEV	Number of children ever born
-	WHYMD1	Reason for first marriage dissolution (divorce or death of husband)
-	WHYMD2	Reason for second marriage dissolution (divorce or death of husband)
-	LAS	Detailed living arrangement at survey (15 categories) [*new*]
-	LIVARR	Living arrangement at survey (4 categories) [*new*]
-	BIRTH1	Number of children during delivery of first child [*new*]
-	BIRTH2	Number of children during delivery of second child
-	BIRTH3	Number of children during delivery of third child
-	BIRTH4	Number of children during delivery of fourth child

Box 3
Modules SPSS Syntax

	Name (*.SPS)	Task
1	1_OG98engWP	Creates labels of variables and values in English Creates the data file NLOG98_F.SAV
2	2_Select variables	Selects a subset of variables from the original file Creates NLOG98_F_2.SAV
3	3_surveymonth	Imputes the survey month Creates a new version of NLOG98_F_2.SAV and file NLOG98_F_cmcint.sav'
	31_Match files INT	Matches two files: NLOG98_F_2.SAV and the file containing one variable (CMCINT) [required when the CMC of month of survey is not re-estimated]
4	4_CMC events	Estimates the CMC of significant events and determines, for each respondent, the number of marriages and cohabitations Creates NLOG98_F_3.SAV
5	5_Marriage sequence	Determines the CMC of first marriage, second marriage and current (third) marriage
6	6_Cohabitation sequence	Determines the CMC of first and subsequent cohabitations
7	7_Leaving home	Estimates the CMC of leaving home
8	8_Remove inconsistencies	Removes inconsistencies in CMCs of marriages Cohabitations and leaving parental home Changes interview month if event after interview
9	9_Birth sequence	Determines singletons, twins, and triplets Changes interview month if birth after interview Creates NLOG98_F_4.sav
10	10_Living arrangement	Living arrangement at time of survey
11	11_Save	Saves the results Creates NLOG98_F_CMC.SAV

b. CMC variables

- CMCINT CMC at interview
- CMCB_OP CMC at birth
- CMCLEAVE CMC at leaving parental home
- CMCMA1 CMC at first marriage
- CMCMA2 CMC at second marriage
- CMCMA3 CMC at third marriage
- CMCE1MA CMC at end of first marriage
- CMCE2MA CMC at end of second marriage
- CMCCO1 CMC at first cohabitation
- CMCE1CO CMC at end first cohabitation
- CMCCO2 CMC at second cohabitation
- CMCE2CO CMC at end second cohabitation
- CMCCO3 CMC at third cohabitation
- CMCE3CO CMC at end third cohabitation
- CMCCO4 CMC at fourth cohabitation
- CMCCO5 CMC at fifth cohabitation
- CMC_K1 CMC at birth of first child
- CMC_K2 CMC at birth of second child
- CMC_K3 CMC at birth of third child
- CMC_K4 CMC at birth of fourth child
- CMC_K5 CMC at birth of fifth child
- CMC_S1 CMC at birth of oldest adopted/foster/stepchild
- CMC_S2 CMC at birth of second adopted/foster/stepchild
- CMC_S3 CMC at birth of third adopted/foster/stepchild

c. Flag variables

- F_SMS Inconsistency between reported marital status at survey date and marital status that is consistent with the marital career and cohabitation sequences
  F_SMS = 1 The reported marital status (BURGS_OP) is inconsistent with the marital status based on the last event in the marital career and the cohabitation career (INTMS).

- FTEST_LH Flag associated with reason for leaving home (see Table 4).

- F_BEWARE Flag indicating that original measurements have been changed (when errors in the data were evident and plausible values could be identified).

- F_K12345 Flag associated with birth to the following women: ID = 1569, 3571, 4098 (see 9_birth sequence)

- FE_GEB CMC birth respondent
- FE_CMA CMC current marriage
- FE_1MA CMC first marriage (if not current)
- FE_1MAEN CMC end first marriage
- FE_2MA CMC second marriage
- FE_2MAEN CMC end second marriage
- FE_COC CMC current cohabitation
- FE_COCM CMC cohabitation before current marriage
- FE_COMA1 CMC cohabitation before first marriage (if not current marriage)
- FE_COMA2 CMC cohabitation before second marriage
- FE_CO1 CMC first cohabitation (if not followed by marriage)
- FE_CO1EN CMC end first cohabitation
- FE_CO2 CMC second cohabitation (if not followed by marriage)
- FE_CO2EN CMC end second cohabitation
- FE_CO3 CMC third cohabitation (if not followed by marriage)
- FE_CO3EN CMC end third cohabitation
- FE_CCO CMC cohabitation contract
- FE_K1 CMC birth oldest child
- FE_K2 CMC birth second child
- FE_K3 CMC birth third child
- FE_K4 CMC birth fourth child
- FE_K5 CMC birth youngest child (not necessarily fifth)
- FE_S1 CMC birth oldest adopted/foster/stepchild
- FE_S2 CMC birth second adopted/foster/stepchild
- FE_S3 CMC birth third adopted/foster/stepchild

- F_CMCMA1 CMC first marriage
  Module: 5_Marriage sequence.sps
  1. Current marriage is first marriage
  2. Current marriage is reported as first marriage, but AANTKEER >
  3. First of two or more marriages

4. Several marriages reported, but CMC available for current marriage only.

- F_CO1 CMC first cohabitation
  Module: 6_Cohabitation sequence.sps
  1. First cohabitation is cohabitation before current marriage
  2. First cohabitation is cohabitation before first marriage
  3. First cohabitation is cohabitation before second marriage
  4. First cohabitation is current cohabitation
  5. First cohabitation is first cohabitation, not followed by a marriage

- F_CO2 CMC second cohabitation
  Module: 6_Cohabitation sequence.sps
  1. Second cohabitation is cohabitation before current marriage
  2. Second cohabitation is cohabitation before first marriage
  3. Second cohabitation is cohabitation before second marriage
  4. Second cohabitation is current cohabitation
  5. Second cohabitation is first cohabitation, not followed by a marriage
  6. Second cohabitation is second cohabitation, not followed by a marriage

- F_CO3 CMC third cohabitation
  Module: 6_Cohabitation sequence.sps
  1. Third cohabitation is cohabitation before current marriage
  2. Third cohabitation is cohabitation before first marriage
  3. Third cohabitation is cohabitation before second marriage
  4. Third cohabitation is current cohabitation
  5. Third cohabitation is first cohabitation, not followed by a marriage
  6. Third cohabitation is second cohabitation, not followed by a marriage
  7. Third cohabitation is third cohabitation, not followed by a marriage

- F_CO4 CMC fourth cohabitation
  Module: 6_Cohabitation sequence.sps
  1. Fourth cohabitation is cohabitation before current marriage
  2. Fourth cohabitation is cohabitation before first marriage
  3. Fourth cohabitation is cohabitation before second marriage
  4. Fourth cohabitation is current cohabitation
  5. Fourth cohabitation is first cohabitation, not followed by a marriage

6. Fourth cohabitation is second cohabitation, not followed by a marriage
7. Fourth cohabitation is third cohabitation, not followed by a marriage

- F_LEAVE CMC: leaving parental home
  Module: 7_Leaving home.sps
  1. Woman leaves home after first cohabitation
  2. Woman leaves home after first marriage
  3. Woman leaves home at onset of cohabitation
  4. Woman leaves home at marriage

- FA_MAR Inconsistencies in marriage
  Module: 8_Remove inconsistencies.sps
  1. Second marriage before end of first marriage or year of end of first marriage missing. CMC second marriage changed to be equal to CMC end first marriage
  2. No end of first marriage is reported, but a second marriage is reported. CMC of end of first marriage is equated to CMC second marriage
  3. No end of second marriage is reported but a third marriage is reported (can be current). CMC of end of second marriage is equated to CMC third marriage.

- FA_CO Inconsistencies in cohabitation
  1. First cohabitation before end of first marriage. CMC end of first marriage equated to CMC of onset of first cohabitation
  2. Second cohabitation before end of first marriage. CMC end of first marriage equated to CMC of onset of second cohabitation
  3. Third cohabitation before end of first marriage. CMC end of first marriage equated to CMC of onset of third cohabitation
  4. First cohabitation before end of second marriage. CMC end of second marriage equated to CMC of onset of first cohabitation
  5. Second cohabitation before end of second marriage. CMC end of second marriage equated to CMC of onset of second cohabitation
  6. Third cohabitation before end of second marriage. CMC end of second marriage equated to CMC of onset of third cohabitation

- FA_S* Inconsistencies in estimated interview date
  a. FA_SLH = 1 Respondent leaves home after survey date. CMCINT is adjusted to be equal to CMCLEAVE
  b. FA_SMA1 =1 First marriage is after survey date. CMCINT is adjusted to be equal to CMCMA1
  c. FA_SME1 = 1 End of first marriage is after survey date. CMCINT is adjusted to be equal to CMCE1MA
  d. FA_SME2 = 1 End of second marriage is after survey date. CMCINT is adjusted to be equal to CMCE2MA
  e. FA_SMA2 = 1 Second marriage after survey date. CMCINT is adjusted to be equal to CMCMA2
  f. FA_SMA3 = 1 Third marriage is after survey date. CMCINT is adjusted to be equal to CMCMA3
  g. FA_SCO1 = 1 First cohabitation is after survey date. CMCINT is adjusted to be equal to CMCCO1
  h. FA_SCO2 = 1 Second cohabitation is after survey date. CMCINT is adjusted to be equal to CMCCO2
  i. FA_SCO3 =1 Third cohabitation is after survey date. CMCINT is adjusted to be equal to CMCCO3
  j. FA_SCE1 = 1 End of first cohabitation is after survey date. CMCINT is adjusted to be equal to CMCE1CO
  k. FA_SCE2 = 1 End of second cohabitation is after survey date. CMCINT is adjusted to be equal to CMCE2CO
  l. FA_SCE3 = 1 End of third cohabitation is after survey date. CMCINT is adjusted to be equal to CMCO3EN
  m. FA_SK1 = 1 Birth of first child is after survey month. CMCINT is adjusted to be equal to CMC_K1.
  n. FA_SK2 = 1 Birth of second child is after survey month. CMCINT is adjusted to be equal to CMC_K2.
  o. FA_SK3 = 1 Birth of third child is after survey month. CMCINT is adjusted to be equal to CMC_K3.
  p. FA_SK4 = 1 Birth of fourth child is after survey month. CMCINT is adjusted to be equal to CMC_K4.
  q. FA_SK5 = 1 Birth of fifth child is after survey month. CMCINT is adjusted to be equal to CMC_K5.

8. CONCLUSIONS AND RECOMMENDATIONS

The Netherlands Family and Fertility Survey 1998 (OG98) is a rich and unique data set. This report describes the conversion of the Public Use Data File of the OG98 into an event history data structure and presents the associated SPSS syntax. The reason for the conversion is that the OG98, like most traditional standardized question list surveys, presents the data in a way that complicates the analysis of event histories (Belli et al., 2001, p. 46). In the OG98, autobiographical information on life events, such as leaving home, cohabitation, marriage, and childbearing, is collected retrospectively. Evidence has shown that memory performance worsens with longer intervals between the occurrence of the events and the retrieval of the information. A consequence is that the sequence of events as well as the dates of events may be misreported. Techniques such as the Life History Calendar have been developed to resolve many of the problems associated with memory loss. The aim of our study was to use the OG98 data as made available by Statistics Netherlands (CBS) and to re-organize the data into a structure that emphasizes the sequence and timing of events in the lives of women. The result is consistent with the information on partnership, marital and fertility careers of Dutch women, as illustrated in Box 1 and Box 2. That data structure is similar to the one proposed by Blossfeld and Rohwer (1995) for the analysis of event histories.

The conversion revealed a number of data problems. The first is incomplete information. Some women reported the occurrence of events but omitted the year of occurrence. Several women did not report the year of birth of at least one of her children. One woman (ID 1569) with five children (AANTLEV) reported the year and month of birth of the youngest child only. In many cases, the month of the event is missing. In these cases, the month is imputed assuming that the event took place in the middle of the year (month 6 = June). We did not use the information on the part of the year in which the event occurred (TRI*). Some women (e.g. ID 815) reported the year and month of first marriage dissolution but did not report the year and month of first marriage.

The other problem is inconsistency. Two sources of inconsistencies are distinguished. The first is a consequence of wrong input data. The second arises when the sequence of events in the life course does not make sense, i.e. when the sequence of events is not plausible. One woman (ID 3571) married

in June 1968 and reported the birth of a first child in July 1974 and the birth of a second child in April 1974. In some cases, it is evident that the error is not due to misreporting but to inaccurate copying. For instance, one woman (ID 2059) has a foster child born in 1872. It is evident that the year of birth should be 1972. Another woman (ID 4646), born in September 1961, left home in April 1961. She indicated that she cohabited before her current marriage (JRTAL), which was in September 1981. From the data on leaving home and marriage, it is clear that the year of leaving home should be 1981. Instead of removing the cases with such anomalies in the data, we decided to examine these cases individually. The reason was that some of the women reporting seemingly inconsistent information have a life course that deviates from what we may expect. For instance, several women start cohabitation or get married while living in the parental home. Initially, we considered those cases as misreporting or anomalies in the data. Later, when we looked at other information provided by these women, we concluded that the situation may be real. When the data error was clear, we changed the figure and designated a flag variable (F_BEWARE=1) to indicate that we made a change in the original data. Flagging the variable allows the individual cases to be easily retrieved. Two women (ID 144 and 999) started cohabitation, married a few months later, and left the parental home several months later. A woman (ID 144) born in 1951, started cohabitation in 1971 and married in 1973. She reported the year and month of leaving home as October 1987 (JRTAL16 and MND16). Should the year read 1967 instead of 1987 or did she really stay in the parental home that long? She did not have to take care of one of her parents since their marriage is intact and both are alive. Another illustration of the complexity of converting the OG98 into an event history file is the fact that several women start cohabiting while married. It is evident from the data that in most cases the onset of cohabitation is related to a divorce. Consider respondent ID 1879, who at time of interview was 39. She left the parental home and married in July 1979. The marriage was dissolved in August 1983 and the reason was divorce. However, she started her first cohabitation in January 1983 that is seven months prior to the divorce. Couldn’t she wait to start living with another man until her husband left? Or did she continue to live with her husband but divorced for financial reasons? She separated from her husband 6 months prior to the survey (VRG25). Is this an anomaly in the data, an isolated case, or a new trend? Who can tell? The interpretation of cases that demonstrate a behaviour that we do not expect as measurement problems might mask signals of social change.

Some information on the sequence and timing of events is difficult to accept. Two examples illustrate the point. First, two women indicated that they remarried before the end of the previous marriage. Second, it seems that one woman (ID171) was married and cohabiting at the same time. She reported two parallel careers in such detail that makes one believe that the living arrangement is real. She reported 5 cohabitations and 1 marriage. The first cohabitation started in August 1973 (CMC 884) and ended in September 1974 (CMC 897). The reason was other than a broken relationship or death of the partner. In the same month, the second cohabitation started. It lasted till November 1975 (CMC 912). The dissolution of the cohabitation was for a reason other than a broken relationship or death of the partner. In 1975, she started a third cohabitation. The month is missing and is imputed to be June; the CMC is therefore 906. This imputation implies that she started the third cohabitation before the end of the second, which may not be realistic. She reports the third cohabitation as one that does not lead to marriage (cmcco3) and that lasts till 1991 (month missing and imputed to be 1098). However, she also reports a cohabitation that does lead to marriage and that starts in the same year as the third one (1975, month missing). She should not have reported so. The imputed CMC is 906. In May 1977, she married (CMC 929) and the marriage lasted till December 1991 (CMC 1104). No reason for marriage dissolution was reported. She separated from her husband 14 months prior to the dissolution that is likely to have been a divorce. During that period, in April 1991 (CMC1096), she started to cohabit with the partner she was living with at the time of the interview. It was her fourth and current cohabitation. This case and many other cases demonstrate the need for a data collection system that pays attention to event histories and the complex life histories that some people experience. By way of illustration, we reported individual cases because they demonstrate best how difficult it may be to distinguish facts from fiction (measurement or reporting errors). The inconsistencies are not limited to isolated cases. For instance, in a number of cases, the marital status at interview, reported by the respondents, is not consistent with the marital status following the last event.

The major conclusions of this study and the recommendations that follow are:

a. The conversion of the OG98 into an event history data structure is feasible but requires attention to individual cases because a considerable number of women in the Netherlands have complex partnership and marital careers. It remains unclear whether the complexity is real or an anomaly of the data.

b. The traditional classification by marital status or living arrangement is inadequate. The concept of state space provides a useful mental frame to introduce structure into the complex types of partnerships. The study demonstrates that a distinction is necessary between marital status and living arrangements. Changes in living arrangements and changes in marital status should be viewed as two parallel careers instead of a single career (partnership career). It is recommended that different careers should be distinguished and that information on each career should be collected and/or recorded separately. At least three careers should be distinguished in fertility and family surveys: the housing career (living arrangement), marital career (marital status), and fertility career. The position in any one of the careers should not be hypothesized based on the position in other careers but should be measured directly. There is some experience with the collection of multiple careers or biographies (e.g. Courgeau and Lelièvre, 1992; Brúckner and Ulrich Mayer, 1998).

c. The event of leaving the parental home must be clearly conceptualized and operationalized.

d. In future surveys, the months and years of life events should be converted into Century Month Code (CMC). Once the dates are recoded into CMCs, inconsistencies in sequence and timing can easily be identified. Other demographic surveys, such as the Demographic and Health Surveys, report the CMCs at events and designate a flag variable to each CMC to indicate whether the month and year was observed or imputed. In the case of imputation, the flag indicates the method of imputation. The OG2003 should adopt a similar approach.

e. The imputation of the CMC at time of interview was necessary because the information was missing in the Public Use File. The date of interview is essential for life history analyses that use hazard models. The reason is that the duration of exposure must be determined accurately, for the rates to be estimated with sufficient accuracy. The imputation of the CMC at interview resulted in a few cases with an event after survey date. The survey date was adjusted accordingly (4 cases). In some cases, the survey date obtained in that way was July 1998 although the OG98 was completed in June 1998.

f. The sequence of cohabitations was particularly difficult to reconstruct. The cohabitations that ended in a marriage are treated differently from cohabitations that do not end in a marriage. A maximum of three cohabitations that do not end in a marriage are considered. We were able to reconstruct sequences of up to five cohabitations. One woman

reported five cohabitations (ID 171). The month of the end of the fourth cohabitation remains unknown, however. It is recommended that all cohabitations should be recorded chronologically. Similarly, the sequence of all marriages and marriage dissolutions should be recorded chronologically. The practice adopted in the Demographic and Health Surveys and the Standard Recode Files of the Fertility and Family Surveys in countries of the ECE region should be considered in future surveys.

g. The questionnaire design and the routing imply several assumptions about the lives of people. For instance, the OG98 assumes that a cohabitation before a marriage is with the future husband. Although the assumption is plausible, it should be made explicit. It is however better to obtain the information directly from the respondent.

h. The OG98 provides the opportunity to study biological children and adopted children, foster children and stepchildren. The month and year of birth of adopted/foster/stepchildren are given. The month and year when the woman adopted the child or when the child started living with her as a stepchild are not given, however. It is recommended that important missing information should be recorded. The sex of the adopted, foster or stepchild should be recorded too.

REFERENCES

BALTES, P.B. and J.R. NESSELROADE (1979), History and rationale of longitudinal research. In: J.R. Nesselroade and P.B. Baltes (eds.), Longitudinal research in the study of behavior and development. Academic Press, New York.

BELLI, R.F., W.L. SHAY and F.P. STAFFORD (2001), Event history calendars and question list surveys. *Public Opinion Quarterly* 65: 45-74.

BLOSSFELD, H.-P., and G. ROHWER (1995), Techniques of event history modeling. New approaches to causal analysis. Lawrence Erlbaum, Mahwah, New Jersey.

BRUECKNER, E. and. K. ULRICH MAYER (1998), Collecting life history data. Experience from the German Life History Study. In: J.Z. Giele and G.H. Elder Jr. (eds.), Methods of life course research. Qualitative and quantitative approaches. Sage Publications, Thousand Oaks, California, pp. 152-181.

COURGEAU, D. and E. LELIÈVRE (1992), Event history analysis in demography. Clarendon Press, Oxford.

DE GRAAF, A. and L. STEENHOF (1999), Relatie- en gezinsvorming van generaties 1945-1979: uitkomsten van het Onderzoek Gezinsvorming 1998 (Partnership and family formation of cohorts 1945-1979: results of the Netherlands Fertility and Family Survey 1998). *Maandstatiek van de Bevolking* 49 (12): 21-36.

DE GRAAF, A. and E. LODEWIJCKX (2000), Technical appendix: Information on the Belgian and Dutch Fertility and Family Surveys. In: J. de Beer and F. Deven (eds.), Diversity in Family Formation: The 2nd Demographic Transition in Belgium and the Netherlands. Kluwer Academic Publishers, Dordrecht, pp.131-139.

ELDER, G.H. Jr. (1999), The life course and aging: some reflections. Distinguished Scholar Lecture, American Sociological Association, August 1999.

FREEDMAN, D., A. THORNTON, D. CAMBURN, D. ALWIN and L. YOUNG-DEMARCO (1988), The life history calendar: a technique for collecting retrospective data. *Sociological Methodology* (18): 37-68.

GARSSEN, J., J. de BEER, P. CUYVERS and A. DE JONG (2001), Samenleven. Nieuwe feiten over relaties en gezinnen. (Cohabitation. New facts about relationships and families). Statistics Netherlands (CBS), Voorburg.

GIELE, J.Z. and G.H. ELDER Jr. (1998), Life course research: development of a field. In: J.Z. Giele and G.H. Elder Jr. (eds.), Methods of life course research. Qualitative and quantitative approaches. Sage Publications, Thousand Oaks, California, pp. 5-27.

KHATUN, M. and F.J. WILLEKENS (2001), The life history calendar. Technical aspects of data analysis using contraceptive calendar of the Bangladesh Demographic and Health Survey. Working Paper 01-2, Population Research Centre, University of Groningen, Groningen.

McDONALD, P. (1996), Demographic life transitions: an alternative theoretical paradigm. *Health Transition Review*, Supplement 6: 385-392.

MILLS, M. (1999), Construction of input data for log-linear models of event histories. Working Paper 99-3, Population Research Centre, University of Groningen, Groningen.

MILLS, M. (2000), The transformation of partnerships. Canada, the Netherlands, and the Russian Federation in the age of modernity. Thela Thesis, Amsterdam.

RYDER, N. (1965), The cohort as a concept in the study of social change. *American Sociological Review* 30(6): 843-861

SCOTT, J. and D. ALWIN (1998), Retrospective versus prospective measurements of life histories in longitudinal research. In: J.Z. Giele and G.H. Elder Jr. (eds.), Methods of life course research. Qualitative and quantitative approaches. Sage Publications, Thousand Oaks, California, pp. 98-127.

TUMA, N. and M. HANNAN (1984), Social dynamics. Models and methods. Academic Press, Orlando.

VERMUNT, J.K. (1997), Log-linear models for event histories. Sage Publications, Thousand Oaks, California.

WILLEKENS, F.J. (2001), Theoretical and technical orientations toward longitudinal research in the social sciences. *Canadian Studies in Population* 28 (2): 189-201 (Special Issue on Longitudinal methodology)

YAMAGUCHI, K. (1991), Event history analysis. Sage Publications, Newbury Park, California.